Integrative studies of 3D genome organization and chromatin structure

Marco Di Stefano, Giacomo Cavalli

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Integrative studies of 3D genome organization and chromatin structure

Marco Di Stefano and Giacomo Cavalli

Abstract

The structural organization of the genome is emerging as a crucial regulator of the cell state, affecting gene transcription, DNA replication, and repair. Over the last twenty years, increasing evidence prompted the development of new experimental techniques to study genome structure. In parallel with the complexity of the novel techniques, computational approaches have become an essential tool in any structural genomics laboratory to analyze and model the data. For biologists to be able to apply the most appropriate modeling approach, it is fundamental to understand the conceptual bases of distinct methods and the insights they can provide. Here, we will discuss recent advances that were possible thanks to 3D genome modeling, discuss their limitations and highlight future perspectives.

Addresses

Institute of Human Genetics, CNRS and University of Montpellier, 34094 Cedex 5, 141 Rue de la Cardonille, 34090, Montpellier, France

Corresponding author: Cavalli, Giacomo (giacomo.cavalli@igh.cnrs.fr) (Di Stefano M.), (Cavalli G.)

Current Opinion in Structural Biology 2022, **77**:102493

This review comes from a themed issue on **Protein Nucleic Acid Interactions**

Edited by **Catherine A. Musselman** and **Eugene Valkov**

For complete overview of the section, please refer the article collection - Protein Nucleic Acid Interactions

Available online 3 November 2022

https://doi.org/10.1016/j.sbi.2022.102493

Keywords

Genome structural organization, 3D genome modeling, Data integration, Top-down methods, Bottom-up methods.

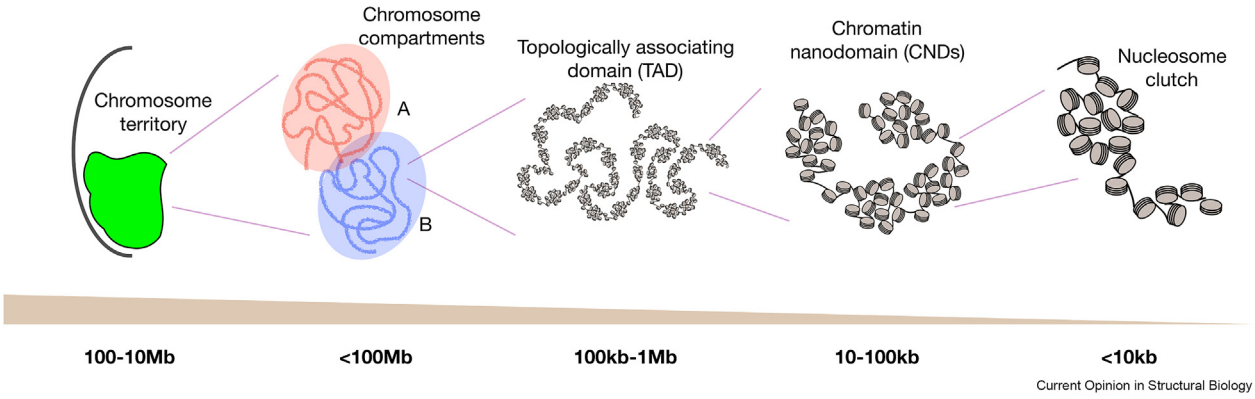
Introduction

In the last two decades, advances in structural genomics revealed that the three-dimensional (3D) genome organization is deeply involved in the regulation of gene expression [1], DNA replication [2], and repair [3]. Two main experimental approaches fostered these discoveries: imaging [4] and chromosome conformation capture (3C-based) techniques [5,6]. Imaging can measure the localization of chromosome loci in the nucleus. These assays can quantify the volume and shape occupied by chromosomal regions or the spatial distance between specific *loci* at various resolutions ranging from several micrometers (μm) to tens of nanometers (nm). 3C-based experiments provide complementary information to imaging. They probe the number of interactions between pairs of chromosome *loci* and typically allow one to measure the contacts of thousands of pairs of regions in a population of millions of nuclei.

These techniques enabled the 3D genome community to characterize multiple layers of the 3D genome organization [7] (Figure 1). Entire chromosomes occupy distinct territories spanning from 1 to a few micrometers [4] with limited intermingling [8] and gene-dependent radial location [9–11]. At the multi-megabase range, chromatin segregates into spatial active (A) and inactive (B) compartments characterized by distinctive GC-content, gene density, and chromatin marks [5,12,13]. At the megabase scale, genomes organize into topologically associating domains (TADs). These are dense self-contacting regions that promote the spatial proximity between genomic *loci* that are distant in the linear genome sequence. Since they may host promoter-enhancer (P-E) contacts, TADs are considered to represent functional units of the genome [14–16]. At the hundred kilobase scale, recent experiments propose an additional layer of organization called chromatin nanodomains (CND) [17]. At the scale of oligo-nucleosomes, a final level of folding is that of nucleosome clutches [18,19] of a size in the scale of a few tens of nanometers.

A crucial aspect of these recent discoveries is the parallel introduction of computational tools for reliable analysis and modeling of experimental data. 3D genome modeling encompasses three distinct, complementary strategies. Top-down approaches use experimental measures as input to generate 3D models representing the data [20]. For example, they can help understand the 3D folding path corresponding to a given 3C-based contact pattern, identifying multi-loci interactions from pair-wise contact information. Bottom-up strategies propose and test first-principle biophysical mechanisms that stem from biological knowledge based on experimental evidence and may constitute general rules of chromatin folding *in vivo* [21]. These approaches use data to test the hypothesized mechanisms. Hybrid

Figure 1



Schematic illustration of chromosome folding inside the nucleus. Chromosomes occupy distinct volumes, named chromosome territories. Chromatin is then segregated into compartments with different epigenetic features. At the mega-base scale, chromatin folds into TADs. Chromosomes form chromatin nanodomains (CNDs) and nucleosome clutches at finer scales.

modeling merges the two previous methods. It integrates experimental data into the models like in top-down approaches and relies on a polymer-based representation of the chromatin fiber as in bottom-up strategies [22].

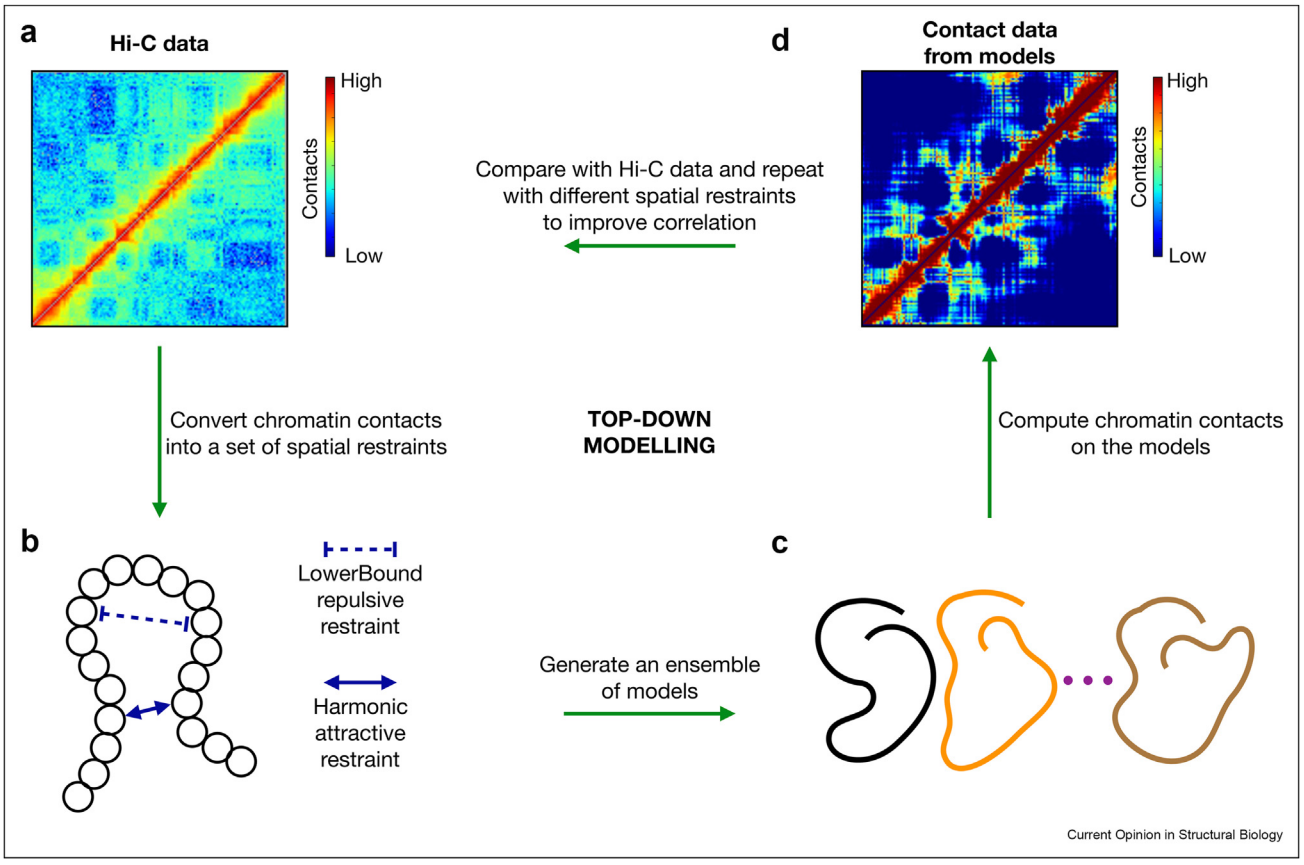
Here, we will present applications of 3D genome modeling studies to showcase how they helped provide novel biological insights otherwise inaccessible to experiments. We will discuss challenges and propose feasible solutions to pursue future developments in 3D genome modeling.

Integrating experimental data into 3D models

The large amount of data generated by imaging and 3C-based techniques prompted the development of data-driven modeling approaches to provide 3D representations of genomic regions of interest and offer valuable insights into how the genome folds in different experimental conditions.

One of the most used data-driven approaches is restraint-based modeling (Figure 2). Its applications integrated data from a plethora of distinct 3C-based techniques, including 5C [23,24], tethered conformation capture [25], 4C [26], Hi-C [27–32], Hi-C time-series [22,28], or Capture-C [33]. However, the methodological steps to generate the models are independent of the input data (Figure 2). Briefly, the chromatin region of interest (Figure 2a) is partitioned into particles, which depending on the application can be the bins of the 3C matrix or entire chromosome domains (TADs). Each particle is represented as a sphere with a diameter proportional to its DNA content. Biophysical models of chromatin compaction, such as the 10 nm or 30 nm fiber models [34], are used to estimate the dimensions of the spheres. The 10 nm-fiber represents chromatin as a chain of nucleosomes connected by linker DNA (20–80 bp) which form repetitive motifs of about 200 bp. The 30 nm fiber model is based on *in vitro* evidence that chromatin assumes a compact solenoidal arrangement in the presence of the linker histone H1 or Mg^{2+} ions [35]. The 30 nm fiber starts, however, to be abandoned since it is generally not seen *in vivo* [36]. Next, the 3C-based contacts are converted into (harmonic) long-range spatial restraints between particle pairs (Figure 2b). The harmonics aim to reproduce the experimental contact pattern in the ensemble of model structures. Hence, genomic regions with a significantly high number of 3C counts, that are most probably close in the 3D space, are restrained with harmonics that force them to be in contact in the model structure. On the other hand, pairs of regions with few 3C counts typically stay far apart, so they are restrained to be separated in the models using lower-bound harmonics (Figure 2B). Next, computer simulations are used to arrange the model particles to satisfy most imposed restraints (Figure 2c). The initial conformation of the model particles is typically a set of disconnected spheres randomly distributed in the simulated space [23]. In some approaches, all the spatial restraints act simultaneously on each model structure, and simulations are repeated on several random initial conformations to resample the set of foldings compatible with the imposed restraints. In other strategies, called population-based [29], only a subset of the restraints is imposed on each initial conformation, so that the entire model's ensemble reproduces the input 3C-based contacts. Both approaches aim to recapitulate the cell-to-cell variability in chromatin structure characterized by single-cell Hi-C and imaging experiments [37,38]. Finally, the ensemble of

Figure 2



Scheme of top-down 3D modeling approaches from Hi-C data. (a) Example of a Hi-C data used as input of data-driven modeling. **(b)** The bins of the Hi-C map are represented as particles of a polymer chain and the contacts are converted into distance restraints. Low numbers of contacts are typically transformed into (lower-bound) repulsive harmonics, and high numbers of contacts into attractive harmonics. **(c)** Numerical simulations are used to organize the particles in space in such a way that the resulting structure satisfies most of the imposed restraints. Due to the fact that it is usually impossible to satisfy all the restraints, which may be also contradictory, repeating the simulations leads to distinct structures, which overall make an ensemble of structures. **(d)** Finally, one computes a contact map from the generated structures and compares it to the input Hi-C data. Iteratively, several sets of restraints are tested to find those which allows recapitulating better the input Hi-C contact pattern. Overall, top-down modeling provides the ensemble of structures that best recapitulates the input chromatin contact pattern.

structures with the best degree of restraint satisfaction (Figure 2d) is considered to represent the most accurate folding describing the input data and are used for downstream analysis.

The biological insights provided by the restraint-based models encompass the comparison of the structural organization in different conditions, including TADs rearrangements before and after progestin treatment in breast cancer cells [31], the folding of the *Hox* clusters in zebrafish and amphioxus [26], active and inactive states in *Arabidopsis thaliana* gene clusters [33], architectural changes during induced-senescence [28] or mouse B-to-PS cells reprogramming [22]. Other studies allowed capturing genome-wide arrangements to study the role of supercoiling in bacteria [24], unveil the functional implications of chromosome positioning in the eukaryotic nucleus [29,32], and analyze the interactions between TADs and nuclear lamina during human adipose stem cells differentiation [30]. For example, in the study by Acemel et al. [26], the authors performed 4C experiments on the *Hox* gene cluster in embryos of zebrafish and amphioxus. 4C measures contacts between a single *locus* (bait) with all the other *loci* genome-wide (*one-vs-all*), hence providing an incomplete picture of the contact patterns. Interestingly, the authors devised a data-driven strategy to obtain structural models in both organisms from 4C data. The models allowed them to obtain virtual *all-vs-all* contact patterns extending the results of the 4C experiments. Using virtual maps, the authors showed that the amphioxus *Hox* cluster folds into a single topological domain that accommodates interactions between *Hox* genes and regulatory elements in the anterior side of the *locus*. In zebrafish, instead, this large region is split into an anterior and a posterior domain. The *Hox* cluster is

located at the boundary and can swing between the two domains. The authors tested this prediction by *ATAC-seq* experiments that allowed them to identify open enhancer-like chromatin regions. This experiment showed that, in fact, the anterior side of the *Hox* gene cluster in amphioxus hosts several ATAC-seq peaks that can be associated with putative distal enhancers. In this case, models were crucial to drive new experiments and to strengthen the conclusion of the study: the bipartite *Hox* domain in vertebrates is an evolutionary novelty that combines pre-existing regulatory contacts in the anterior *Hox* domain with new ones in the posterior one.

Alternative data-driven approaches are based on the hypothesis that pairs of particles interact via short-range (Lennard-Jones) attractive potentials [39–42]. The initial state of the chromatin regions is an already connected random (self-avoiding walk) polymer. The short-range interactions make a substantial difference from the data-driven approaches discussed above because in this case the contacts are established only if the two particles come sufficiently close in space during the simulation, or if their interaction is mediated by binding factors. For example, in [39] Bianco et al. used a machine learning procedure called polymer-based recursive statistical inference method (PRISMR) to derive the optimal binding domains and domain-interaction strengths from the Hi-C interaction patterns of the *HoxD* locus in mouse embryonic stem cells (mESC) and cortical neuron cells (CNC) [43]. Using the String & Binders Switch (SBS) model, the authors obtained structural models at 5 kb resolution, revealing broader inter-TADs interactions in CNCs than in ESCs, leading to the formation of higher-order structures (metaTADs).

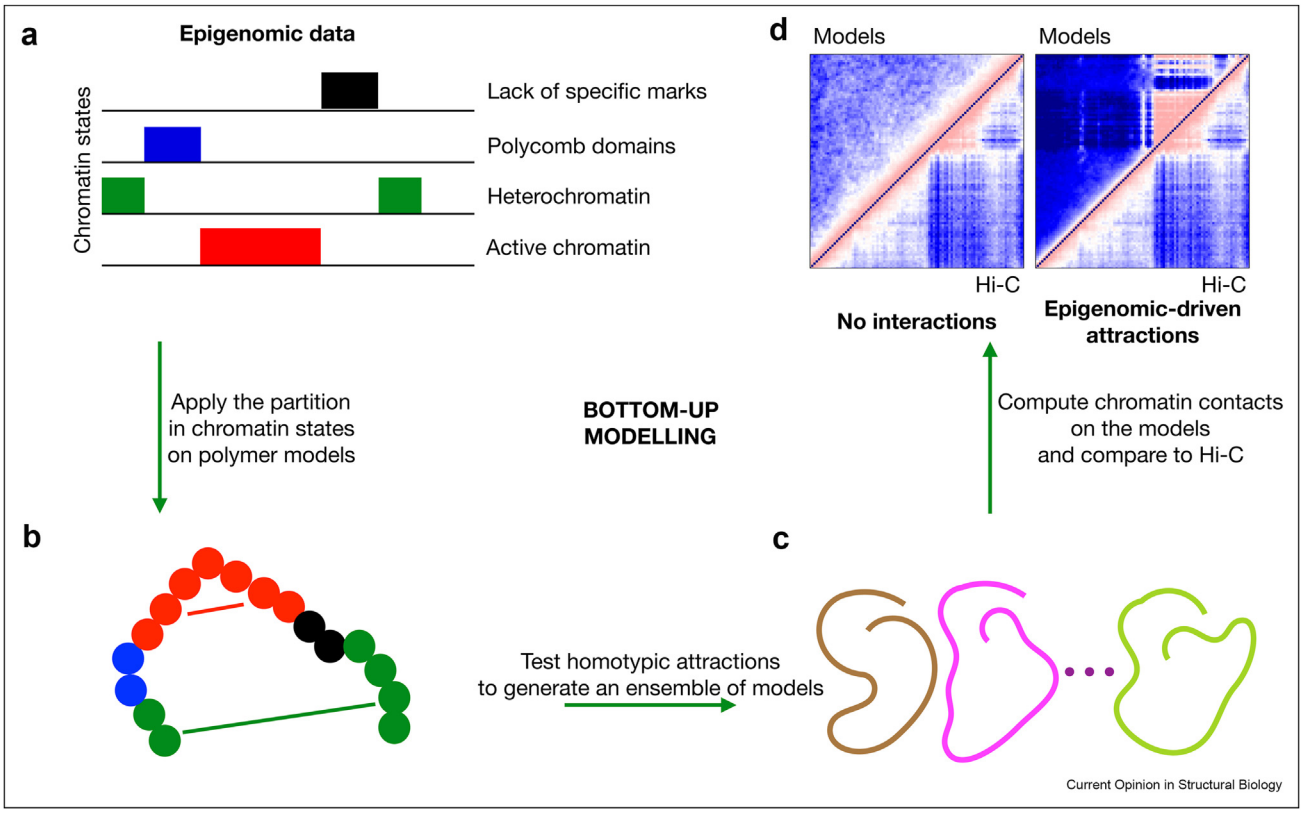
Complementary to 3C-based methods, recent imaging techniques [44–48] are fostering novel data-driven genome modeling approaches [44,45]. For example, in [44], the authors developed a novel experimental technique combining super-resolution microscopy methods (OligoSTORM and OligoDNA-PAINT). They studied a region of 8 Mb of human chromosome 19 and visualized structures at different scales from intra-TADs to entire compartments in single cells. Like Cryo-EM techniques for proteins, this experimental technique provides a cloud of points that defines a volume where the tagged chromatin regions are most probably located. To take advantage of this data, the authors developed Integrative Modeling of Genomic Regions (IMGR) to rearrange an ensemble of structures obtained from Hi-C data at 10 kb resolution on the cloud by optimal fitting. Using these structures informed on the imaging data, the authors provided evidence that chromosome compartments are physical structures with high cell-to-cell variability and variable degree of mutual overlap. Interestingly, the structural arrangements of the chromosome regions, accessible only via 3D modeling, showed significantly higher variability between homologous chromosomes than between chromosomes in different cells, suggesting a parental-dependent chromosome folding, whose functional implications ought to be further explored.

Testing biophysical mechanisms via 3D genome modeling

Complementary to top-down approaches, bottom-up modeling moves from formulating hypotheses on which biophysical processes may regulate the 3D genome folding. First, data analysis leads to the inference of general physical principles or testable mechanistic rules that may shape the 3D genome. Next, these rules are implemented in numerical simulations of a polymer chain describing chromatin. At this stage, the methods rely mainly on computations to test many different values of the model's parameters but can also integrate experimental data for the simulation setup. For example, the nuclear volume and shape can limit the dimension of the simulated environment, the physical properties of chromatin (*e.g.*, chromatin thickness or bending rigidity) informs the polymer model, and nuclear DNA and volumetric density define the crowding of the simulated system [49]. Some approaches [49,50] also use experimental data to pin down the optimal models' parameters. Finally, the obtained structures are tested against experimental data. Overall, the outcomes of bottom-up approaches are the 3D models of the region of interest, the identification of the biophysical processes, which likely contribute to shape the genome organization, and the quantitative estimates of the forces regulating these processes.

One example of bottom-up approaches aimed to unveil the biophysical mechanisms regulating genome compartmentalization. These modeling approaches moved from two main observations. First, epigenomic features, including histone modifications and transcription factor binding, led to the partition of the chromosome sequence into chromatin states with distinctive transcription-regulation activities [51,52] (Figure 3a). Second, Hi-C maps revealed that this one-dimensional partition of the genome reflects a three-dimensional organization into spatial domains because regions of the same chromatin state tend to contact each other more often than different ones [49,52]. These observations suggest that phase-separation between chromatin of distinct types may drive 3D genome organization [16,53]. For example, in *Drosophila melanogaster*, Jost et al. [53] devised a bottom-up modeling strategy to test the hypothesis that attractive short-range interactions between chromatin domains of the same type may reproduce the 3D genome folding of several regions (Figure 3b–c). The polymer models could predict compartments with good accuracy using four chromatin states, including active, Polycomb

Figure 3



Scheme of bottom-up epigenetics-driven modeling. (a) The starting point is the analysis of chromatin immunoprecipitation (ChIP-seq) data to identify the binding of DNA- and histone-associated proteins. This analysis allows classifying chromatin in distinct states depending on epigenomic features. For instance, in recent work on the plant *A. thaliana* [49], one can identify active chromatin, heterochromatin, polycomb repressive chromatin, and chromatin that lacks specific histone marks (null). **(b)** Next, chromatin states are mapped on polymer models of chromosomes or chromosome regions. **(c)** Numerical simulations are used to test homotypic attractive interactions between chromatin states. Varying the strength of the attractions allows generating several ensembles of models. **(d)** One computes contact maps from each ensemble of models and compares them with Hi-C interaction data. For instance, we show two cases without (*left*) and with (*right*) epigenomic-driven attractions: the interactions induce the formation of chromosome compartments, which appear as a checkerboard pattern in the contact map. The final results of this bottom-up procedure are the model parameters and the structural models associated with the highest correlation to Hi-C.

repressed, heterochromatin, and null (Figure 3d). Further works explored how the physics of phase-separation leads to the formation of compartments in different species, their maintenance over time, or their dynamical changes along the cell cycle or developments [49,54,55].

Intriguingly, concurrent mechanisms of chromosome compartmentalization may involve RNA molecules [56—58]. In particular, Lu et al. found significant correlations between the partition of chromosomes in active (A) and repressive (B) compartments in mouse and human cells and the detection of the DNA sequence of the genome of Long and Short Interspersed Nuclear Elements (LINE and SINE). Specifically LINE are enriched in B and SINE in A chromosome compartments. This observation suggests that a mesh of RNA molecules that are the transcripts of the repetitive DNA sequences leads to the formation and stabilization of chromatin compartments [56]. In Farabella et al., the authors proposed and tested via bottom-up modeling a role for long non-coding (lncRNA)/DNA triplexes to regulate the coating of repressive domains in specific accumulation sites leading to compartment segregation [58].

Another example of how bottom-up modeling helped propose mechanisms of 3D genome organization is the loop-extrusion model [59], which can lead to the formation of TADs in interphase [50,60]. Briefly, SMC complexes [61] land on chromatin, start extruding loops until they unbind, bump into each other, or encounter CTCF complexes in convergent orientation, which act as barriers. These predictive models fostered experimental and modeling studies to test loop-extrusion *in vitro* and *in vivo* [62—64] and to extend its applicability to mitotic chromosome compaction and segregation by

condensins [65–67], DNA repair [3], the folding of bacterial DNA [68], and transcription [69].

Current bottlenecks of 3D genome modeling

3D genome modeling has already brought great insight into genome folding principles, but it has to overcome limitations.

In the 3D genomics community, we miss a molecular understanding of how P-E interactions work in three-dimensional space. There is no consensus on fundamental quantitative aspects, such as *what is the typical spatial distance between a promoter and an enhancer when the gene is active or inactive? How long should this spatial distance be maintained in time in order to maintain transcriptional activity? How frequently should P-E pairs interact to activate a gene, and how long should these interactions last?* A correlated question regards the involvement of proteins or regulatory factors: *How many proteins are needed to activate or repress a gene?* To address these questions, we need experiments to simultaneously probe the localizations of several genomic *loci* and the spatial positions of proteins and transcription machinery (*e.g.*, Polymerase components) at high resolution. In parallel, 3D modeling should study long DNA filaments connecting regulatory DNA elements and their promoters at high spatial (atomic) resolution (P-E regulation may take place over hundreds of kilo-bases), as well as the regulatory factors involved (*e.g.*, proteins and histone modifications).

Currently, the predictive power of 3D modeling approaches is focused mainly on the folding of the genome, with a limited understanding of how it is related to the functional state of the cell. In particular, it has been shown that disruption of TAD architecture can lead to ectopic P–E contacts and thus gene misexpression, contributing to developmental defects [70], cancer [71], and autism susceptibility [72]. At this stage, 3D modeling cannot integrate effectively and systematically genomic factors and their manipulations in the models. These data may include transcription factor binding and histone modification or DNA sequence modifications, like deletions, duplications or insertions. Integrating these data and understanding their effect on the 3D genome organization may lead to predict, for instance, how TAD rearrangements (creation or disruption) might switch genes on or off, leading to changes of cell states.

Additionally, the community needs to widen the spectrum of experimental techniques to tackle open questions and to improve the resolution of the existing ones. The diversification of approaches can enable us to study, for example, the 3D organization of repetitive sequences that may have originated from chromosome duplications or insertions of transposable elements. For instance, a recent work used Atomic Force Microscopy (AFM) [73] to characterize the folding of centromeres in human cells. The DNA sequences of these regions are rich in a unique type of DNA tandem repeats, called α-satellites, and for this reason, are inaccessible to techniques based on genome sequencing, such as 3C. This approach might be extended to study in general how repetitive regions fold. 3D modeling should contribute by developing predictive high-resolution models to study how the DNA sequence affects chromosome folding and ultimately gene expression.

Conclusions

The discoveries provided by 3D genome modeling in recent years make us confident that we will be able to address the challenges we are currently facing in the field in the next few years. In the following, we suggest possible actions for the structural genomics community to deepen the understanding of the 3D genome and its interplay with gene function.

Recent high-resolution imaging [17–19,46,47] and micro-C [74,75] studies are starting to unveil the nucleosome-scale structure of the chromatin fiber. They show that chromatin is a heterogeneous fiber formed by groups (*clutches*) of 2–20 nucleosomes and a thickness varying in the 9–24 nm range. Modeling strategies may integrate this data to define a more realistic representation of the chromatin fiber in 3D computational studies, challenging the 10 nm or 30 nm chromatin models [76]. In parallel, high-throughput imaging experiments [13,77] provide useful single-cell information on the localization of thousands of specific loci in the genome. 3D modeling studies may use this data to characterize the P-E interactions in single cells or to validate and challenge their predictions on genome folding.

For a significant breakthrough in understanding the genome structure–function relationship, it is crucial to improve the computational performance of 3D modeling methods. For example, the community may achieve this objective by developing highly-parallel codes that can scale well on several CPUs or GPUs or by generating on-lattice models which allow for much faster simulations and parameter screening [78]. Another desirable venue is to design multi-scale models that can efficiently simulate different chromatin scales within the same system. These approaches would allow a currently missing comprehensive knowledge to fill all the gaps within and between the 3D genome organizational layers (Figure 1). This knowledge would help us connect variations in the DNA sequence directly to changes in P-E contacts and, finally, to the functional state of the cell.

Another important aspect is to encourage the development and sharing of free, open-source software and

generated models following the Findability, Accessibility, Interoperability, and Reusability (FAIR) principles, as it is for experimental data. This practice would facilitate data re-analysis and algorithm benchmarking [79,80] and foster collaborations between experimental and computational research groups, which brought a lot of insight in recent years [81,82]. From this synergy, we can expect a huge boost in the development of 3D genome modeling.

Funding

This work was supported by the European Research Council (Advanced Grant 3DEpi, under grant agreement number 788972); the European Union (CHROMDESIGN Project, under the Marie Skłodowska-Curie grant agreement number 813327); the Fondation pour la Recherche Médicale (DEI20151234396); the MSDAVENIR foundation (project GENE-IGH); the Centre National pour la Recherche Scientifique; the Agence Nationale de la Recherche (E-RARE project "IMPACT" and project ANR-18-CE15-0010); and the French National Cancer Institute (INCaPLBIO18-362).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

We apologize to those whose work could not be discussed due to space limitations. We thank all the members of the Cavalli lab for their extensive discussions and great insights.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Sexton T, Cavalli G: **The role of chromosome domains in shaping the functional genome**. *Cell* 2015, **160**:1049–1059.
2. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, *et al.*: **Topologically associating domains are stable units of replication-timing regulation**. *Nature* 2014, **515**:402–405.
3. Arnould C, Rocher V, Finoux A-L, Clouaire T, Li K, Zhou F, *et al.*: **Loop extrusion as a mechanism for formation of DNA damage repair foci**. *Nature* 2021, **590**:660–665.
4. Cremer T, Cremer C: **Chromosome territories, nuclear architecture and gene regulation in mammalian cells**. *Nat Rev Genet* 2001, **2**:292–301.
5. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, *et al.*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **326**:289–293.
6. Dekker J, Rippe K, Dekker M, Kleckner N: **Capturing chromosome conformation**. *Science* 2002, **295**:1306–1311.
7. Jerkovic I, Cavalli G: **Understanding 3D genome organization by multidisciplinary methods**. *Nat Rev Mol Cell Biol* 2021, **22**:511–528.
8. Branco MR, Pombo A: **Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations**. *PLoS Biol* 2006, **4**:e138.
9. Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA: **The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells**. *Hum Mol Genet* 2001, **10**:211–219.
10. Tanabe H, Habermann FA, Solovei I, Cremer M, Cremer T: **Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications**. *Mutat Res* 2002, **504**:37–45.
11. Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, *et al.*: **Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes**. *PLoS Biol* 2005, **3**:e157.
12. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping**. *Cell* 2014, **159**:1665–1680.
13. Su J-H, Zheng P, Kinrot SS, Bintu B, Zhuang X: **Genome-scale imaging of the 3D organization and transcriptional activity of chromatin**. *Cell* 2020, **182**:1641–1659. e26.
14. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, *et al.*: **Topological domains in mammalian genomes identified by analysis of chromatin interactions**. *Nature* 2012:376–380, https://doi.org/10.1038/nature11082.
15. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, *et al.*: **Spatial partitioning of the regulatory landscape of the X-inactivation centre**. *Nature* 2012, **485**:381–385.
16. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, *et al.*: **Three-dimensional folding and functional organization principles of the Drosophila genome**. *Cell* 2012, **148**:458–472.
17. Szabo Q, Donjon A, Jerković I, Papadopoulos GL, Cheutin T, Bonev B, *et al.*: **Regulation of single-cell genome organization into TADs and chromatin nanodomains**. *Nat Genet* 2020, **52**:1151–1157.
18. Ricci MA, Manzo C, García-Parajo MF, Lakadamyali M, Cosma MP: **Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo**. *Cell* 2015, **160**:1145–1158.
19. Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH, O'Shea CC: **ChromEMT: visualizing 3D chromatin structure and compaction in interphase and mitotic cells**. *Science* 2017, **357**, https://doi.org/10.1126/science.aag0025.
20. Serra F, Di Stefano M, Spill YG, Cuartero Y, Goodstadt M, Baù D, *et al.*: **Restraint-based three-dimensional modeling of genomes and genomic domains**. *FEBS Lett* 2015, **589**:2987–2995.
21. Chiang M, Brackley CA, Marenduzzo D, Gilbert N: **Predicting genome organisation and function with mechanistic modelling**. *Trends Genet* 2021, https://doi.org/10.1016/j.tig.2021.11.001.
22. Di Stefano M, Stadhouders R, Farabella I, Castillo D, Serra F, Graf T, *et al.*: **Transcriptional activation during cell reprogramming correlates with the formation of 3D open chromatin hubs**. *Nat Commun* 2020, **11**:2564.
23. Baù D, Sanyal A, Lajoie BR, Capriotti E, Byron M, Lawrence JB, *et al.*: **The three-dimensional folding of the α-globin gene domain reveals formation of chromatin globules**. *Nat Struct Mol Biol* 2011, **18**:107–114.
24. Umbarger MA, Toro E, Wright MA, Porreca GJ, Baù D, Hong S-H, *et al.*: **The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation**. *Mol Cell* 2011, **44**:252–264.

25. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L: **Genome architectures revealed by tethered chromosome conformation capture and population-based modeling**. *Nat Biotechnol* 2011, **30**:90–98.

26. Acemel RD, Tena JJ, Irastorza-Azcarate I, Marlétaz F, Gómez-Marín C, de la Calle-Mustienes E, *et al.*: **A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation**. *Nat Genet* 2016, **48**:336–341.

27. Serra F, Baù D, Goodstadt M, Castillo D, Filion GJ, Marti-Renom MA: **Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors**. *PLoS Comput Biol* 2017, **13**, e1005665.

28. ** Sati S, Bonev B, Szabo Q, Jost D, Bensadoun P, Serra F, *et al.*: **4D genome rewiring during oncogene-induced and replicative senescence**. *Mol Cell* 2020, **78**:522–538. e9.
The authors merged cutting-edge 3D Genomics experiments with both top-down and bottom-up modeling strategies to characterize genome structure and dynamics during oncogene-induced senescence (OIS) in human cells. Models suggested that chromatin detachment from the nuclear lamina and weakening of interactions between Senescence-associated Heterochromatin Domains (SAHD) contribute to regulating 3D Genome reorganization in IOS.

29. Tjong H, Li W, Kalhor R, Dai C, Hao S, Gong K, *et al.*: **Population-based 3D genome structure analysis reveals driving forces in spatial genome organization**. *Proc Natl Acad Sci U S A* 2016, **113**:E1663–E1672.

30. Paulsen J, Sekelja M, Oldenburg AR, Barateau A, Briand N, Delbarre E, *et al.*: **Chrom3D: three-dimensional genome modeling from Hi-C and nuclear lamin-genome contacts**. *Genome Biol* 2017, https://doi.org/10.1186/s13059-016-1146-2.

31. Le Dily F, Baù D, Pohl A, Vicent GP, Serra F, Soronellas D, *et al.*: **Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation**. *Genes Dev* 2014, **28**:2151–2162.

32. Li Q, Tjong H, Li X, Gong K, Zhou XJ, Chiolo I, *et al.*: **The three-dimensional genome organization of Drosophila melanogaster through data integration**. *Genome Biol* 2017, https://doi.org/10.1186/s13059-017-1264-5.

33. Nützmann H-W, Doerr D, Ramírez-Colmenero A, Sotelo-Fonseca JE, Wegel E, Di Stefano M, *et al.*: **Active and repressed biosynthetic gene clusters have spatially distinct chromosome states**. *Proc Natl Acad Sci U S A* 2020, **117**:13800–13809.

34. Zhu P, Li G: **Structural insights of nucleosome and the 30-nm chromatin fiber**. *Curr Opin Struct Biol* 2016:106–115, https://doi.org/10.1016/j.sbi.2016.01.013.

35. Finch JT, Klug A: **Solenoidal model for superstructure in chromatin**. *Proc Natl Acad Sci U S A* 1976, **73**:1897–1901.

36. Maeshima K, Imai R, Tamura S, Nozaki T: **Chromatin as dynamic 10-nm fibers**. *Chromosoma* 2014, **123**:225.

37. Cattoni DI, Cardozo Gizzi AM, Georgieva M, Di Stefano M, Valeri A, Chamousset D, *et al.*: **Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions**. *Nat Commun* 2017, **8**:1753.

38. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, *et al.*: **Single-cell Hi-C reveals cell-to-cell variability in chromosome structure**. *Nature* 2013, **502**:59–64.

39. * Bianco S, Annunziatella C, Andrey G, Chiariello AM, Esposito A, Fiorillo L, *et al.*: **Modeling single-molecule conformations of the HoxD region in mouse embryonic stem and cortical neuronal cells**. *Cell Rep* 2019, **28**:1574–1583.e4.
The authors introduced a machine learning procedure (PRISMR) to infer the polymer model that optimally describes chromatin contact data and applied a polymer-based modeling strategy (SBS) to obtain the corresponding 3D structure. The work is an application of these tools to study the structure-function relationship in the *HoxD* cluster in mouse embryonic stem cells and cortical neurons.

40. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, *et al.*: **Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription**. *Cell* 2014, **157**:950–963.

41. Zenk F, Zhan Y, Kos P, Löser E, Atinbayeva N, Schächtle M, *et al.*: **HP1 drives de novo 3D genome reorganization in early Drosophila embryos**. *Nature* 2021, **593**:289–293.

42. Barbieri M, Chotalia M, Fraser J, Lavitas L-M, Dostie J, Pombo A, *et al.*: **Complexity of chromatin folding is captured by the strings and binders switch model**. *Proc Natl Acad Sci USA* 2012:16173–16178, https://doi.org/10.1073/pnas.1204799109.

43. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, *et al.*: **Multiscale 3D genome rewiring during mouse neural development**. *Cell* 2017, **171**:557–572.e24.

44. ** Nir G, Farabella I, Pérez Estrada C, Ebeling CG, Beliveau BJ, Sasaki HM, *et al.*: **Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling**. *PLoS Genet* 2018, **14**, e1007872.
This work introduced a new super-resolution imaging technique based on OligoSTORM and OligoDNA-PAINT and a new top-down modeling method, called IMGR, to extract from imaging and Hi-C data single-cell models of the chromatin folding. Focussing on 8 megabases of human chromosome 19, the authors found that, despite sizable cell-to-cell variability, chromosome compartments are structured physical objects in single cells.

45. Nguyen HQ, Chattoraj S, Castillo D, Nguyen SC, Nir G, Lioutas A, *et al.*: **3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing**. *Nat Methods* 2020, https://doi.org/10.1038/s41592-020-0890-0.

46. Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, Boettiger AN: **Visualizing DNA folding and RNA in embryos at single-cell resolution**. *Nature* 2019, **568**:49–54.

47. Bintu B, Mateo LJ, Su J-H, Sinnott-Armstrong NA, Parker M, Kinrot S, *et al.*: **Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells**. *Science* 2018, **362**, https://doi.org/10.1126/science.aau1783.

48. Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, *et al.*: **Super-resolution imaging reveals distinct chromatin folding for different epigenetic states**. *Nature* 2016, **529**:418–422.

49. * Di Stefano M, Nützmann H-W, Marti-Renom MA, Jost D: **Polymer modelling unveils the roles of heterochromatin and nucleolar organizing regions in shaping 3D genome organization in Arabidopsis thaliana**. *Nucleic Acids Res* 2021, **49**:1840–1858.
Using a bottom-up modeling approach, this work studies the principles regulating genome folding in the *A. thaliana* nucleus. Only four chromatin states, including constitutive and facultative heterochromatin, active chromatin, and nucleolar organizing regions, can recapitulate chromosome genome-wide at 3 kilo-bases resolution. Models reveal an essential organizational role for nucleolar and constitutive heterochromatin regions.

50. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA: **Formation of chromosomal domains by loop extrusion**. *Cell Rep* 2016, **15**:2038–2049.

51. Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, *et al.*: **Systematic protein location mapping reveals five principal chromatin types in Drosophila cells**. *Cell* 2010, **143**:212–224.

52. * Di Pierro M, Cheng RR, Lieberman Aiden E, Wolynes PG, Onuchic JN: **De novo prediction of human chromosome structures: epigenetic marking patterns encode genome architecture**. *Proc Natl Acad Sci U S A* 2017, **114**:12126–12131.
Using energy landscape theory and neural network-based machine learning, the authors predicted the 3D structure of the human genome at 50 kb resolution from epigenetic marks. 3D models are validated by comparisons with chromosome capture and imaging data.

53. Jost D, Carrivain P, Cavalli G, Vaillant C: **Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains**. *Nucleic Acids Res* 2014, **42**:9553–9561.

54. Brackley CA, Johnson J, Kelly S, Cook PR, Marenduzzo D: **Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains**. *Nucleic Acids Res* 2016, **44**:3503–3512.

55. * Buckle A, Brackley CA, Boyle S, Marenduzzo D, Gilbert N: **Polymer simulations of heteromorphic chromatin predict the 3D folding of complex genomic loci**. *Mol Cell* 2018, **72**:786–797. e11.

The authors introduced HiP-HoP, a polymer model to predict 3D genome structure from epigenetic and protein binding data. The simulations suggest substantial cell-to-cell structural variability and different mechanisms of enhancer action.

56. Lu JY, Chang L, Li T, Wang T, Yin Y, Zhan G, *et al.*: **Homotypic clustering of L1 and B1/Alu repeats compartmentalizes the 3D genome**. *Cell Res* 2021, **31**:613–630.

57. Quinodoz SA, Jachowicz JW, Bhat P, Ollikainen N, Banerjee AK, Goronzy IN, *et al.*: **RNA promotes the formation of spatial compartments in the nucleus**. *Cell* 2021, **184**:5775–5790. e30.

58. Farabella I, Di Stefano M, Soler-Vila P, Marti-Marimon M, Marti-Renom MA: **Three-dimensional genome organization via triplex-forming RNAs**. *Nat Struct Mol Biol* 2021, **28**:945–954.

59. Alipour E, Marko JF: **Self-organization of domain structures by DNA-loop-extruding enzymes**. *Nucleic Acids Res* 2012, **40**: 11202–11212.

60. Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, *et al.*: **Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes**. *Proc Natl Acad Sci U S A* 2015, **112**:E6456–E6465.

61. Banigan EJ, Mirny LA: **Loop extrusion: theory meets single-molecule experiments**. *Curr Opin Cell Biol* 2020, **64**:124–138.

62. Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, *et al.*: **Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization**. *Cell* 2017, **169**:930–944. e22.

63. Rao SSP, Huang S-C, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-R, *et al.*: **Cohesin loss eliminates all loop domains**. *Cell* 2017, **171**:305–320.e24.

64. ** Gabriele M, Brandão HB, Grosse-Holz S, Jha A, Dailey GM, Cattoglio C, *et al.*: **Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging**. *Science* 2022, **376**:496–501.

The authors studied the folding of the CTCF/cohesin-mediated loop hosting the *Fbn2* gene in mESC using micro-C, imaging, and bottom-up polymer modeling. They inferred that the *Fbn2* loop is predominantly (97% of the time) partially extruded during the cell cycle and assumes a looped state only rarely.

65. Gibcus JH, Samejima K, Goloborodko A, Samejima I, Naumova N, Nuebler J, *et al.*: **A pathway for mitotic chromosome formation**. *Science* 2018, **359**, https://doi.org/10.1126/science.aao6135.

66. Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, *et al.*: **Organization of the mitotic chromosome**. *Science* 2013, **342**:948–953.

67. Goloborodko A, Imakaev MV, Marko JF, Mirny L: **Compaction and segregation of sister chromatids via active loop extrusion**. *Elife* 2016, **5**, https://doi.org/10.7554/eLife.14864.

68. Brandão HB, Ren Z, Karaboja X, Mirny LA, Wang X: **DNA-loop-extruding SMC complexes can traverse one another in vivo**. *Nat Struct Mol Biol* 2021, **28**:642–651.

69. Brandão HB, Paul P, van den Berg AA, Rudner DZ, Wang X, Mirny LA: **RNA polymerases as moving barriers to condensin loop extrusion**. *Proc Natl Acad Sci U S A* 2019, **116**: 20489–20499.

70. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, *et al.*: **Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions**. *Cell* 2015, **161**:1012–1025.

71. Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, *et al.*: **Insulator dysfunction and oncogene activation in IDH mutant gliomas**. *Nature* 2016: 110–114, https://doi.org/10.1038/nature16490.

72. Lybaek H, Robson M, de Leeuw N, Hehir-Kwa JY, Jeffries A, Haukanes BI, *et al.*: **LRFN5 locus structure is associated with autism and influenced by the sex of the individual and locus conversions**. *Autism Res* 2022, https://doi.org/10.1002/aur.2677.

73. Chardon F, Japaridze A, Witt H, Velikovsky L, Chakraborty C, Wilhelm T, *et al.*: **CENP-B-mediated DNA loops regulate activity and stability of human centromeres**. *Mol Cell* 2022, **82**: 1751–1767. e8.

74. Krietenstein N, Abraham S, Venev SV, Abdennur N, Gibcus J, Hsieh T-HS, *et al.*: **Ultrastructural details of mammalian chromosome architecture**. *Mol Cell* 2020, **78**:554–565. e7.

75. Hsieh T-HS, Cattoglio C, Slobodyanyuk E, Hansen AS, Rando OJ, Tjian R, *et al.*: **Resolving the 3D landscape of transcription-linked mammalian chromatin folding**. *Mol Cell* 2020, **78**:539–553. e8.

76. Fussner E, Ching RW, Bazett-Jones DP: **Living without 30nm chromatin fibers**. *Trends Biochem Sci* 2011, **36**:1–6.

77. Takei Y, Zheng S, Yun J, Shah S, Pierson N, White J, *et al.*: **Single-cell nuclear architecture across cell types in the mouse brain**. *Science* 2021, **374**:586–594.

78. Ghosh SK, Jost D: **How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes**. *PLoS Comput Biol* 2018, **14**, e1006159.

79. Forcato M, Nicoletti C, Pal K, Livi CM, Ferrari F, Bicciato S: **Comparison of computational methods for Hi-C data analysis**. *Nat Methods* 2017, **14**:679–685.

80. Zufferey M, Tavernari D, Ciriello G: **Methods for the analysis of topologically associating domains (TADs)**. *Methods Mol Biol* 2022, **2301**:39–59.

81. Marti-Renom MA, Almouzni G, Bickmore WA, Bystricky K, Cavalli G, Fraser P, *et al.*: **Challenges and guidelines toward 4D nucleome data and model standards**. *Nat Genet* 2018, **50**:1352–1358.

82. Dekker J, Belmont AS, Guttman M, Leshyk VO, Lis JT, Lomvardas S, *et al.*: **The 4D nucleome project**. *Nature* 2017, **549**:219–226.