Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Network models of chromatin structure

Vera Pancaldi*

Increasing numbers of datasets and experimental assays that capture the organization of chromatin inside the nucleus warrant an effort to develop tools to visualize and analyze these structures. Alongside polymer physics or constraint-based modeling, network theory approaches to describe 3D epigenome organization have gained in popularity. Representing genomic regions as nodes in a network enables visualization of 1D epigenomics datasets in the context of chromatin structure maps, while network theory metrics can be used to describe 3D epigenome organization and dynamics. In this review, we summarize the most salient applications of network theory to the study of chromatin contact maps, demonstrating its potential in revealing epigenomic patterns and relating them to cellular phenotypes.

Address
CRCT, Université de Toulouse, Inserm, CNRS, Université Toulouse III-Paul Sabatier, Centre de Recherches en Cancérologie de Toulouse, Toulouse, France

Corresponding author: Pancaldi, Vera (vera.pancaldi@inserm.fr)
* Twitter account: @verapancaldi

Current Opinion in Genetics & Development 2023, **80**:102051

This review comes from a themed issue on **Genome Architecture and Expression**

For complete overview of the section, please refer to the article collection, "Genome Architecture and Expression (2023)"

Edited by **Giacomo Cavalli** and **Job Dekker**

Available online 26 May 2023

https://doi.org/10.1016/j.gde.2023.102051

Introduction

In the last twenty years, we have started producing maps of the structure of the genome inside the nucleus, through chromatin capture experiments and other complementary techniques. Our ability to chart out chromatin structure is revolutionizing our mental models of the genome. For years, we have been representing the genome as a line, prompted by the availability of extremely useful genome browsers, which we used as visualization tools for different types of datasets (gene definitions, chromatin mark peaks, replication origin activity, etc.). Diverse experiments have undeniably shown that DNA is intertwined at various levels in complex 3D structures. Moreover, a variety of RNA and proteins are associated with chromatin, constituting what has been recently termed the nucleome, an important entity that connects the genome (intended as purely genome sequence) to gene regulation and cellular phenotypes [1].

The transition between single-cell microscopy-based assays to chromosome conformation capture technologies has produced an explosion in the quantity and diversity of chromatin structure datasets available. The popularization of next-generation sequencing has enabled the efficient mapping of genome-wide chromatin contacts by conformation capture, for example, through Hi-C [2], or of interactions mediated by specific chromatin-bound factors [3], painting the first pictures of 3D genome structure. Further developments have led to new experimental approaches that lead to even more detailed descriptions of the 3D genome [4–12]. Each experimental technique gives us a partial but useful view on chromatin organization, so comparison and benchmark of these methods are necessary [13,14].

Independent of the different assays that can be used to detect 3D chromatin contacts, we can distinguish between at least 3 types of chromatin structure maps that can be built from these datasets: 1) real spatial descriptions of chromatin inside single cells, through advanced labeling and microscopy approaches, 2) spatial models of the chromatin produced by polymer physics-based or constraint-based simulations, starting from experimental chromatin contact data, and 3) abstract topological representations of chromatin contacts through network models, derived directly from the experimental data. In this review, we will focus specifically on network theory approaches to visualize, interpret, and compare chromatin structure datasets.

Visualizing chromatin as a network

A network is composed of a set of nodes connected by edges (see Box 1) and it is a useful representation to model systems in which different elements (represented as nodes) are related to each other (as shown by edges linking two nodes) either because the two elements interact, or because their properties are somehow correlated. For example, people can be nodes of a social network based on interactions (two people physically meet each other) or because they have similar behavior (they buy similar products). In biology, networks of

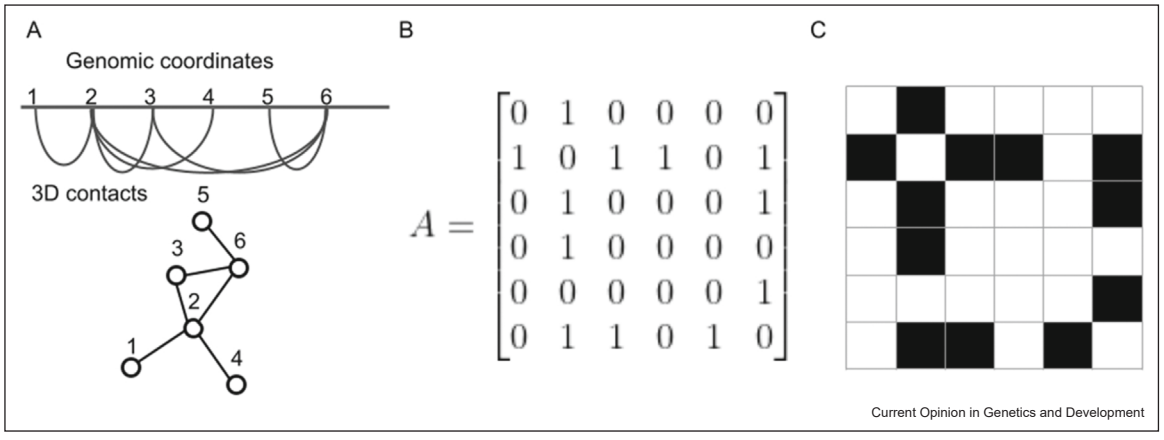
Box 1 Network theory basics.

A network is a collection of objects/elements (nodes, represented by circles in the diagram) connected by links if they interact or if they are related in any way. An example of how genomic contacts can be represented as networks is shown in Figure B1. Networks of N nodes can also be represented as square matrices (NxN) in which a 1, a 0, or a value (weight) at each position of the matrix represents the existence or strength of a connection joining two specific nodes in the network (an adjacency matrix).

The field of network theory is very old, but the relatively recent addition of statistical mechanics concepts and improved availability of computational power in the last two decades has generated an unprecedented interest and number of applications across fields as varied as social science, neurology, economics, and food science, to cite but a few.

The igraph package [67] and NetworkX [68] offer numerous algorithms for calculation of the most common network statistics (see Table 1 for some examples) as well as network visualization. Networks can be represented in 2D using specific algorithms that assign node positions based on user-defined requirements. The most common representations aim to put nodes that are connected close to each other and optimize the layout for visibility (avoiding node overlap). Widely used network visualization tools that allow interactive exploration of the nodes as well as basic statistics calculations are Cytoscape [69] and Gephi [70], while MuxViz [71] can be used to visualize and investigate multilayer networks and multiplexes that are representations in which the same elements in a specific system can be connected according to different criteria, which are represented by different networks shown as layers.

Figure B1



Different representations of genomic contacts using networks. **(a)** 3D contacts between chromatin regions can be shown as arcs on linear genome browsers or as edges (links) connecting nodes that represent the chromatin regions (fragments or genes or other) forming a network. **(b)** A network can be represented as an adjacency matrix, in which element A_{ij} is 1 if there is contact between nodes i and j. **(c)** The adjacency matrix can also be represented as a heatmap and this is the typical way in which Hi-C data are represented. In this case, each chromatin region corresponds to a 'bin' spanning a specific genome length and the number of contacts between each pair of bins can be shown with color intensity (see Figure 1a on the right).

Table 1

Glossary of network-related terms [72,73].

Density (network)	Number of edges over possible number of edges
Connected component	Subset of nodes that can be reached one from the other
Module/community	Set of nodes that show many interconnections with each other (different measures can be defined)
Clique	Set of nodes that have all possible edges between them
Diameter (network)	Length of the longest path between 2 nodes
K-core	Largest subnetwork in which each node has degree > k. Coreness is k when the node belongs to k-core but not (k + 1)-core
Degree centrality (node)	Number of connections (edges)
Betweenness centrality (node/edge)	Number of shortest paths going via the node/edge
Closeness centrality (node)	Average number of steps separating each network node to the node in consideration
Clustering coefficient/transitivity (node)	Proportion of neighbors of the node that are connected over total possible number
Assortativity (feature)	Correlation coefficient describing to what extent nodes with one feature tend to connect with other nodes with the same feature [38]

proteins interacting with each other or of genes with correlated gene expression have been studied for decades and are behind ontologies and functional enrichment, which have driven the transition from small-scale wet-lab biology to genome-wide and omics approaches.

We can represent chromatin as a network by considering nodes to represent chromatin regions/fragments, while edges joining two nodes indicate 3D proximity or contact between the corresponding genomic regions. One of the most popular techniques to map these contacts is Hi-C, which uses the principle of chromosome conformation capture, involving cross-linking of chromatin with formaldehyde, digesting it, and religating it to preserve information about which sequence fragments were close in 3D space, but on a global genome-wide scale. Since 2010, very soon after the first HiC datasets had been produced, models of chromatin structure as a network of chromatin fragments have led to many interesting results.

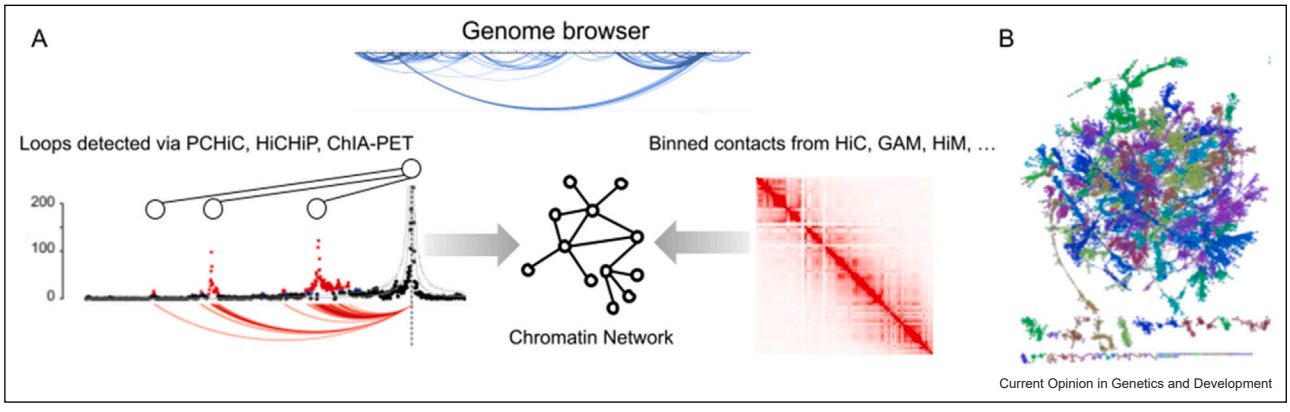
Since genome-wide chromatin structure datasets produced by Hi-C started being represented as matrices, drawn as heatmaps, showing essentially the number of contacts between all possible pairs of chromatin regions, they were quickly interpreted as adjacency matrices (see Box 1) of a chromatin contact network [15–19]. In this case, the number of contacts between specific regions could be used as weight (strength) of the interactions between nodes that represent these regions in the network. The main limitation of the early Hi-C datasets was a lack of resolution and the fact that most of the datasets would describe non-coding regions of the genome (the most abundant overall), making it hard to associate structure patterns to genes and biological function.

Since then, an increasing interest in network analyses was justified by improvements in the quality of the datasets available [20]. At present, there are several web servers that allow visualizing chromatin as networks [21–24]. Before the appearance of two 3D genome browsers [25,26] in 2022, all other genome browsers still represented the genome as a linear object, or as a triangular/square heatmap of contacts [27]. The existence of these 3D network browsers facilitates the analysis of specific chromatin loci and helps to explore distal enhancers that regulate specific genes (Figure 1).

From network topology to function

Beyond the visualization of proximity relationships, there are several advantages in using networks to represent chromatin. The first one is that patterns in the epigenome can be studied at the statistical level, not focussing on specific loci but observing general organizational principles, building on decades of work in the field of network theory. With the development of alternative chromosome conformation capture methods, such as Chromatin Interaction Analysis by Paired-End Tag (ChIA-PET) and capture Hi-C and especially Promoter Capture Hi-C, which focussed on contacts involving a subset of specific genomic regions (either bound by the same chromatin factor or overlapping gene

Figure 1



From linear genome visualizations to chromatin networks. **(a)** Schematic showing how different chromosome conformation datasets can be used to extract chromatin networks. On the top, a typical view of chromatin interaction on a typical genome browser. On the left, we show a typical output of statistical analysis of Capture Hi-C data, with Y axis showing the number of ditags (pairs of paired-end reads spanning a chromatin contact), showing in red the regions that show significant contacts with the rightmost peak in black, which is used as a viewpoint, which is also joined by red arcs below. Significant interactions are shown as edges between nodes. On the right, a schematic of a Hi-C contact matrix heatmap, showing on X and Y a chromosome or chromosome region, in which red signifies a high number of contacts were observed between the two regions. Both types of data can be represented as chromatin networks. In one case, the nodes represent chromatin fragments (produced as part of the Capture Hi-C protocol), in the case of Hi-C, the nodes represent bins in which the genome is subdivided. **(b)** mESC combined Promoter Capture Hi-C and Hi-C chromatin network, displayed using default layout in Cytoscape. Node colors denote chromosomes.

promoters, respectively), it became easier to study the contacts' topology in relation to functional properties at a global genome-wide level. By functional properties, we refer to characteristics of a system, which can affect its behavior and response to its environment, which, translated into biological terms, corresponds to cellular phenotypes. Using several network theory measures applied to chromatin networks in which long-range (>10MB) interactions were preserved, Sandhu et al. discovered that genes were organized in modules related to their functional importance. They also identified a so-called rich club, meaning a set of highly connected genes with roles in important cellular processes that were also connected to each other, while observing that less-connected nodes had functions in more specific processes or developmental stages and were enriched in genetic mutations [28].

Several other papers analyzed chromatin network topology with network theory tools. An application of Markov models on Hi-C-derived chromatin interaction networks highlighted their hierarchical structure [29]. The first pattern that was identified in chromosome capture data was the presence of regions with a lot of interactions within the region and fewer contacts with other regions, creating squares in Hi-C heatmaps that were denominated topologically associating domains (TADs), which have been the focus of a large body of literature ever since they were defined [30]. Given the almost identical definition of network modules (set of nodes displaying more interactions with each other than with other nodes), this analogy led to a TAD detection algorithm [31], while the internal organization of some TAD network modules in a core-periphery structure was found to be biologically and medically relevant [32]. A similar approach to studying spatial aspects of chromatin organization has led to the identification of cliques in the chromatin network, sets of nodes that are fully interconnected with each other [33].

Using chromatin networks to integrate 1D epigenomic datasets in a 3D context

One of the most powerful approaches in biology to extract biological hypotheses from different experimental assays involves data integration. For example, if we have discovered that specific genomics regions are enriched in peaks of a particular transcription factor, we can study which other factors also bind in similar regions and deduce that potentially the two factors are needed for a specific process. An analysis of integration of chromatin marks has led to the definition of different chromatin states [34] and successively chromatin colors [35], which help to define a set of rules that epigenomes tend to obey. Similarly, we might be interested in integrating datasets describing genomic regions (the same definitions of chromatin states for example) with the 3D structure of chromatin, to identify patterns in how regions with specific chromatin states might be organized inside the nucleus. This type of analysis is greatly facilitated by the network analogy.

As is the case in other applications of network theory, nodes can be characterized by feature vectors, enabling the projection of a variety of datasets on the chromatin network (Figure 2**a**). Chromatin fragment features can allow us to establish 'classes' of nodes and analysis of the topology of the different classes has led to several insights. Integrating histone modification data from Chromatin ImmunoPrecipitation sequenecing (ChIP-seq) in three different human cell lines, Thibodeau et al. identified the importance of broad domains and super-enhancers as hubs (highly connected nodes) in the chromatin interaction network and proved that the topology around regulatory elements can predict the functional importance of regulated genes associated with them [36].

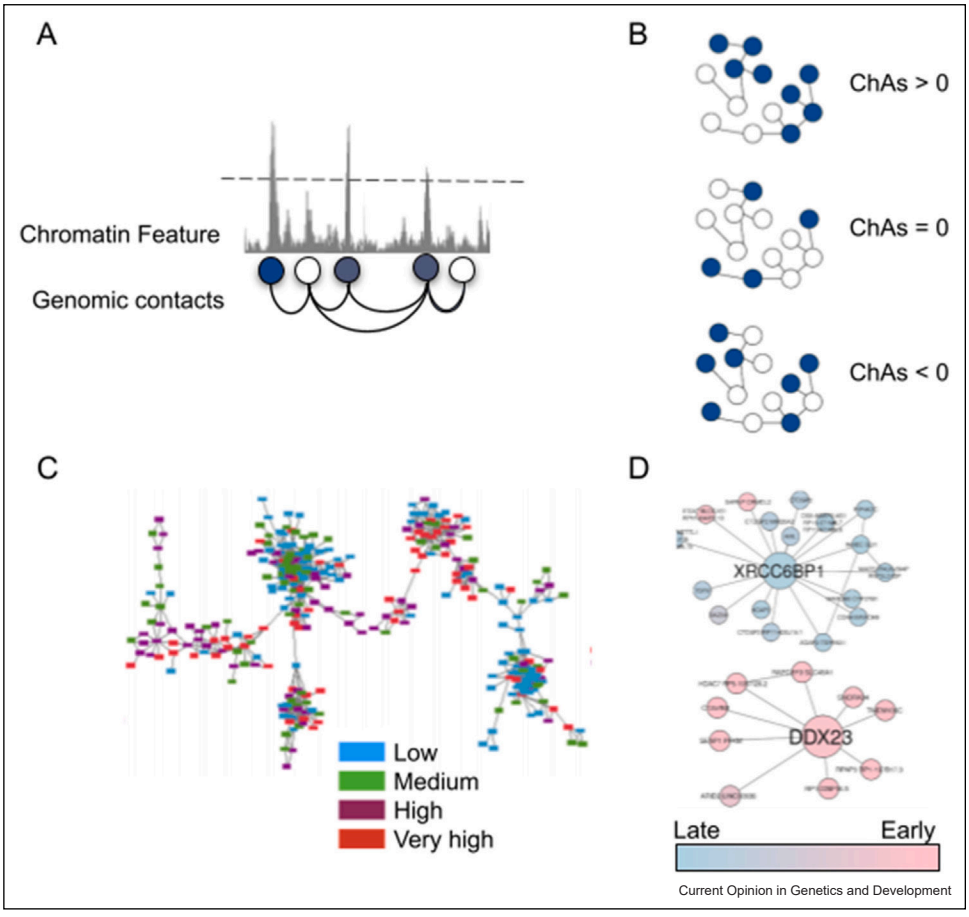
Exploiting a collection of over 80 epigenomic features mapped in mouse embryonic stem cells (mESCs), we introduced the concept of chromatin assortativity (ChAs) to identify whether chromatin fragments with specific features tend to preferentially interact with each other in 3D (Figure 2**b**) [37].

Assortativity, homophily, or assortative mixing is a measure of correlation between values of features on nodes connected by edges in a network, which ranges from – 1 to 1, and whose value expresses whether nodes with similar features tend to be preferentially connected in the network [38]. When considering chromatin networks, ChAs measure whether fragments with a specific value of a feature form clusters, where the features can be any value assigned to a chromatin fragment, including presence of histone modifications, methylation or expression levels, or even attributes of genes in the corresponding chromatin region.

ChAs were applied to the set of features projected on an mESC promoter-capture Hi-C dataset, identifying polycomb group proteins as central players in the maintenance of the interaction core, while we identified biologically relevant differences in ChAs of specific forms of RNA Pol II (paused, initiating, and elongating) in promoter–enhancer subnetworks [37].

ChAs was also applied to analyze monocytes, neutrophils, and T cells' promoter-centered networks to study the relationship between 3D structure and gene expression (Figure 2**c**), DNA methylation, as well as the most important histone marks. The ChAseR package and the GARDEN-NET web server provide the community with tools to calculate assortativity of any desired feature on any chromatin network [22].

Figure 2



ChAs integrates 1D epigenomic datasets onto chromatin networks. **(a)** Schematic representing a quantitative chromatin feature (could represent chromatin accessibility, binding of a Transcription Factor (TF), or replication) projected onto nodes of a chromatin contact network. Node color represents the presence of the chromatin feature on the fragment of chromatin represented by the node, in this case, the node is blue if the chromatin feature reads are above a certain threshold. **(b)** Overview of possible values of ChAs of the feature mapped onto the chromatin network. When blue nodes are preferentially interacting with other blue nodes, ChAs will be positive (top); when bue nodes are interacting with blue or white nodes indifferently, ChAs will be zero; when blue nodes are always interacting with white nodes, ChAs will be negative, the feature represented by the blue color is said to be disassortative. The concept of ChAs is better defined as a correlation for continuous features, but the schematic uses a binary variable for illustration purposes. **(c)** Promoter–promoter network extracted from Promoter Capture Hi-C on monocytes showing gene expression as color of the nodes. **(d)** A concrete example of assortativity of replication timing on the human monocyte Promoter Capture Hi-C network as represented in Genome ARchitecture DNA Epigenome and Nucleome-Network Exploration Tool (GARDEN-NET) [22], with node color representing replication timing.

We also reported positive and very high assortativity of replication timing (Figure 2**d**), generalizing previous observations that TADs represent domains of consistent replication timing [22]. Finally, assortativity of replication origin efficiency across distances even longer than TAD sizes suggests a strong link between replication and 3D chromatin structure at multiple scales [39].

ChAs was also used to detect the localized nature of contacts affected by STAG2 knockdown in bladder cancer cell lines and to identify the typical range of distances spanned by such altered contacts [40]. Finally, ChAs of several transcription factors was used to assess the performance of bioinformatic tools to identify significant contacts from promoter-capture Hi-C datasets [41].

Using network frameworks to relate chromatin structure and gene expression

Some controversies still exist over the strength and direction of the connection between 3D genome structure and gene expression. As soon as Hi-C datasets became available, it was reported that co-expressed genes were

often found close or in 3D proximity of each other [15,42]. TADs were soon identified as domains of coordinated gene regulation, although the existence of coregulated domains distinct from TADs was also reported [43]. Researchers have often considered gene regulation as one of the main purposes behind chromatin structure organization [44]. Recent experiments that were designed to thoroughly test this hypothesis have provided a yet unclear picture. On one side, it is still unclear whether interactions detected by chromosome capture methods really capture physical contacts or proximity, on the other side, simultaneous measurements of chromatin structure features, such as TADs, and live-transcription detection do not find a particularly strong mechanistic link between the two [10,45,46].

However, the reason for the absence of clear patterns could well be our incomplete understanding of how the various regulatory elements, proximal and distal, and specific loops that cannot be captured at the level of large-scale features such as TADs, jointly control expression programs. Some evidence that expression can be related to chromatin structure was provided by network analyses, in terms of assortativity of gene expression levels, which was confirmed in promoter-centered networks for monocytes, neutrophils, and T cells [22]. On the other hand, regulatory regions are also shown to form hubs in these networks. For example, integrating Formaldehyde-assisted isolation of regulatory elements (FAIRE-seq)-detected open-chromatin regions with Hi-C data, thousands of hubs of enhancer and promoter contacts were identified [47]. Interestingly, applying a sensitive technique for mapping interactions specifically involving the MYC gene [48], an important oncogene with high relevance in multiple cancer types, it was suggested that the gene is dynamically establishing contacts with different enhancers [49]. Through network analysis, the authors propose that the MYC promoter itself is scanning the different enhancers, consistent with our own [37] and others' results [50] suggesting that RNA Polymerase II (RNAPII)- mediated interactions would be date-like (with different partners at any one point) rather than party-like (joining all partners at all times). Methods borrowed from the field of network science, in some cases already applied to the study of protein interactomes, can reveal organizing principles in chromatin contact maps and mechanisms involved in gene regulation.

Finally, we can exploit very powerful tools to compare and classify networks using graph neural network approaches. Representing the genome as a network calls for specific deep learning frameworks. GraphReg [51] is a recent tool based on Graph Attention Networks that can predict enhancer activities and gene expression, integrating a variety of features including TF motifs, 1D epigenomic datasets such as chromatin accessibility or histone modifications, with chromatin long-range loops detected by Hi-C-derived techniques. Machine learning has also been used to predict enhancer–promoter 3D contacts from RNA-seq and Assay for Transposase-Accessible Chromatin with high-throughput sequencing (ATAseq) in samples for which obtaining chromatin structure data would be challenging, such as cancer patient samples [52]. We expect more deep learning approaches, especially those building on the network representation of chromatin structure, to appear in the near future and reveal the intricate connection between genome structure and function.

Comparing networks and capturing alterations in their structure

Maps of protein interactions allow us to comprehend the relationships between different pathways and even predict the function of uncharacterized proteins. Studying protein interactomes led scientists to wonder about how interaction networks could have grown through evolution, leading to theories about gene duplication and subfunctionalization. Similarly, looking at chromatin structure networks makes us wonder how the 3D genome evolved through the transition from unicellular to multicellular organisms. Interestingly, Yan et al. discovered that genomic fragments of the *Arabidopsis thaliana* genome that interact in 3D have similar levels of polymorphisms and evolution rates [53]. The finding that most features associated with mutation rates display ChAs suggests an interplay between mutational processes and chromatin structure. Indeed, the assortativity of gene variability and plasticity [54], which are connected to evolutionary processes [55,56], strengthens the evidence for a strong connection between 3D genomic localization and evolution (Raynal et al. in prep.).

A recent paper also proposes a network-based model of chromatin folding, with an analogy to explosive percolation, the phenomenon according to which the successive random addition of contacts between network nodes suddenly generates global connectivity. The authors model the progressive folding of chromatin regions into loops, with different models of loop extrusion, recapitulating the emergence of phase-separated chromatin regions [57]. It remains to be seen whether these models could tell us more about the evolution of the 3D genome. On the other hand, network concepts were applied by Chovanec et al. to describe the alterations in chromatin structure observed in human cells transitioning from naive to primed pluripotent cell states [21] and similar approaches highlighted drastic changes in network topology between normal B cells and malignant counterparts in Chronic Lymphocytic Leukemia (CLL) patients [58].

Chromatin networks as a step toward multilevel phenotypic regulation

Networks are ideally suited to represent chromatin structure datasets produced by different experimental assays. In addition, nested networks of networks, in which nodes can represent entire networks, can be used to visualize the genome in an inherently hierarchical way, thus allowing the integration of disparate information about 3D contacts, without many constraints on the definition of what a network node should correspond to (a gene, a chromatin fragment, an entire TAD [39], and potentially an entire chromosome). This makes networks ideal for descriptions of multiscale and multilevel systems, such as the genome. Using simulations of polymer models, Brackley et al. were able to predict that spatial patterns in 3D interactions between promoters and regulatory regions lead naturally to the emergence of small-world gene-regulatory networks [59], which are networks in which the average distance between any two nodes is shorter than would be expected at random. Also, in this spirit, the APRIL framework introduced by Wang et al. expands lists of disease genes by considering networks of 3D contacts to include long-range gene-regulatory interactions [60]. Interestingly, analyses of chromatin interaction networks in neurons and other cells in the brain have helped to elucidate the genetic basis of brain disorders [61,62] and have also unveiled potential connections between psychiatric and metabolic diseases [63].

A formal way to represent multiple networks interacting is through the use of multiplexes, mathematical objects that represent a set of elements that can interact in different ways, represented as a set of networks stacked one on top of the other, or multilayer networks, in which each network layer can contain different nodes [64]. Despite the fact that genes, mRNAs, and proteins are undoubtedly different biological entities in terms of cellular components, we can think about functional roles of genes, their transcribed mRNAs, and their translated proteins as a single thing. Thus, representing chromatin as a network has facilitated the integration of epigenomic interactions in multilevel biological multiplexes and multilayers [65,66]. Chiliński et al. combine different networks of genes, proteins, functional terms (gene ontology), as well as metabolites to trace relationships and information flow between mutations, Single Nucleotide Polymorphisms (SNPs), and Expression Quantitative Trait Loci (eQTLs), all the way to phenotypes.

Conclusions

The availability of detailed maps of the organization of chromatin inside the nucleus has undoubtedly revolutionized our understanding of the subtle connection between epigenomes and phenotypes. We are progressively discovering that projecting all genomic datasets on a 1D line, an imaginary straightened-out chromosome, was very useful but quite misleading, as it downplayed the role of the physical localization of genome regions in the space inside the nucleus.

Just like specific projections of the world map distort the relative sizes between countries, forcing our understanding of chromatin structure into one dimension through traditional Hi-C heatmaps and linear browsers has affected our ability to comprehend some of the genome's organizational principles. Fortunately, old and new techniques that force us to preserve spatial relationships between genomic regions explicitly, albeit with their own set of limitations, have helped us to develop new, richer visualization and analysis frameworks. Constraint-based and polymer physics-based models of chromatin structure are extremely useful to rescue the three-dimensional character of the epigenome, but the increasing number of applications of this framework suggests that network representations are a useful and complementary approach.

To summarize some of the advantages of this formalism, we highlights the following points: chromatin networks are extremely efficient to build, starting from any kind of chromatin structure datasets, and integrate with additional 1D datasets or information; the flexibility in deciding what a node represents provides the potential to consider hierarchical structures, while mathematical multiplex frameworks simplify the task of integrating epigenome-level regulation with additional regulatory cellular processes. Finally, we anticipate that the development of graph neural network approaches to describe the epigenome will further increase the interest and value of this representation.

Data Availability

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

Declaration of Competing Interest

None.

Acknowledgements

This work was supported by Fondation Toulouse Cancer Santé and Pierre Fabre Research Institute as part of the Chair of bioinformatics in oncology of the CRCT.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Dekker J, *et al.*: **The 4D nucleome project**. *Nature* 2017, **549**:219-226.

2. Erez L-A, *et al*.: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **326**:289-293.

3. Fullwood MJ, *et al*.: **An oestrogen-receptor-α-bound human chromatin interactome**. *Nature* 2009, **462**:58-64.

4. Mumbach MR, *et al*.: **HiChIP: efficient and sensitive analysis of protein-directed genome architecture**. *Nat Methods* 2016, **13**:919-922.

5. Hsieh T-HS, *et al*.: **Mapping nucleosome resolution chromosome folding in yeast by micro-C**. *Cell* 2015, **162**:108-119.

6. Downes DJ, *et al*.: **Capture-C: a modular and flexible approach for high-resolution chromosome conformation capture**. *Nat Protoc* 2022, **17**:445-475.

7. Mifsud B, *et al*.: **Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C**. *Nat Genet* 2015, **47**:598-606.

8. Beliveau BJ, Apostolopoulos N, Wu C: **Visualizing genomes with Oligopaint FISH probes**. *Curr Protoc Mol Biol* 2014, **105**, https://doi.org/10.1002/0471142727.mb1423s105 (Unit 14.23) PMID: 24510436; PMCID: PMC3928790.

9. Cardozo Gizzi AM, *et al*.: **Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms**. *Mol Cell* 2019, **74**:212-222 e5.

10. Espinola SM, *et al*.: **Cis-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early Drosophila development**. *Nat Genet* 2021, **53**:477-486.

11. Takei Y, *et al*.: **Integrated spatial genomics reveals global architecture of single nuclei**. *Nature* 2021, **590**:344-350.

12. Quinodoz SA, *et al*.: **RNA promotes the formation of spatial compartments in the nucleus**. *Cell* 2021, **184**:5775-5790.e30.

13. Fiorillo L, *et al*.: **Comparison of the Hi-C, GAM and SPRITE methods using polymer models of chromatin**. *Nat Methods* 2021, **18**:482-490.

14. Shi G, Thirumalai D: **Conformational heterogeneity in human interphase chromosome organization reconciles the FISH and Hi-C paradox**. *Nat Commun* (1) 2019, **10**:3894, https://doi.org/10.1038/s41467-019-11897-0

15. Babaei S, *et al*.: **Hi-C chromatin interaction networks predict co-expression in the mouse cortex**. *PLoS Comput Biol* 2015, **11**:e1004221.

16. •• Botta M, Haider S, Leung IXY, Lio P, Mozziconacci J: **Intra-and inter-chromosomal interactions correlate with CTCF binding genome wide**. *Mol Syst Biol* 2010, **6**:426, https://doi.org/10.1038/msb.2010.79 PMID: 21045820; PMCID: PMC3010120.
This is to our knowledge the first paper that uses a network formalism to represent chromatin structure datasets.

17. Boulos RE, Arneodo A, Jensen P, Audit B: **Revealing long-range interconnected hubs in human chromatin interaction data using graph theory**. *Phys Rev Lett* 2013, **111**:118102.

18. Boulos RE, *et al*.: **From the chromatin interaction network to the organization of the human genome into replication N/U-domains**. *New J Phys* 2014, **16**:115014.

19. Morlot J-B, Mozziconacci J, Lesne A: **Network concepts for analyzing 3D genome structure from chromosomal contact maps**. *EPJ Nonlinear Biomed Phys* 2016, **4**:2.

20. Bonev B, *et al*.: **Multiscale 3D genome rewiring during mouse neural development**. *Cell* 2017, **171**:557-572 e24.

21. •• Chovanec P, *et al*.: **Widespread reorganisation of pluripotent factor binding and gene regulatory interactions between human pluripotent states**. *Nat Commun* 2021, **12**:2098.
This paper uses network representations and analyses to demonstrate rearrangements in chromatin structure, describing multi-scale organization and communities in these networks.

22. •• Madrid-Mencía M, Raineri E, Cao TBN, Pancaldi V: **Using GARDEN-NET and ChAseR to explore human haematopoietic 3D chromatin interaction networks**. *Nucleic Acids Res* 2020, **48**:4066-4080.
This paper introduces tools for constructing chromatin networks and analyzing them within a dedicated R package, as well as a web-tools to visualize chromatin networks and measure ChAs.

23. Ramirez RN, Bedirian K, Gray SM, Diallo A: **DNA Rchitect: an R based visualizer for network analysis of chromatin interaction data**. *Bioinformatics* 2020, **36**:644-646.

24. Thibodeau A, *et al*.: **QuIN: a web server for querying and visualizing chromatin interaction networks**. *PLoS Comput Biol* 2016, **12**:e1004809.

25. Li D, Harrison JK, Purushotham D, Wang T: **Exploring genomic data coupled with 3D chromatin structures using the WashU Epigenome Browser**. *Nat Methods* 2022, **19**:909-910.

26. Zhu X, *et al*.: **Nucleome Browser: an integrative and multimodal data navigation platform for 4D Nucleome**. *Nat Methods* 2022, **19**:911-913.

27. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, Li D, Choudhary MNK, Li Y, Hu M, Hardison R, Wang T, Yue F: **The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions**. *Genome Biol* (1) 2018, **19**:151, https://doi.org/10.1186/s13059-018-1519-9 PMID: 30286773; PMCID: PMC6172833.

28. Sandhu KS, *et al*.: **Large-scale functional organization of long-range chromatin interaction networks**. *Cell Rep* 2012, **2**:1207-1219.

29. Tan ZW, Guarnera E, Berezovsky IN: **Exploring chromatin hierarchical organization via Markov State Modelling**. *PLoS Comput Biol* 2018, **14**:e1006686.

30. Sexton T, *et al*.: **Three-dimensional folding and functional organization principles of the Drosophila genome**. *Cell* 2012, **148**:458-472.

31. • Norton HK, *et al*.: **Detecting hierarchical genome folding with network modularity**. *Nat Methods* 2018, **15**:119-122.
This paper proposes a TAD detecting algorithm based on networks theory interpretation of TADs as modules in the chromatin network.

32. •• Huang H, *et al*.: **A subset of topologically associating domains fold into mesoscale core-periphery networks**. *Sci Rep* 2019, **9**:9526.
This paper exploits the network concepts of core and periphery to relate chromatin TAD topology to presence of variants associated with diseases.

33. Paulsen J, *et al*.: **Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation**. *Nat Genet* 2019, **51**:835-843.

34. Ernst J, Kellis M: **ChromHMM: automating chromatin-state discovery and characterization**. *Nat Methods* 2012, **9**:215-216.

35. Filion GJ, *et al*.: **Systematic protein location mapping reveals five principal chromatin types in Drosophila cells**. *Cell* 2010,1-13, https://doi.org/10.1016/j.cell.2010.09.009

36. Thibodeau A, Márquez EJ, Shin D-G, Vera-Licona P, Ucar D: **Chromatin interaction networks revealed unique connectivity patterns of broad H3K4me3 domains and super enhancers in 3D chromatin**. *Sci Rep* 2017, **7**:14466.

37. Pancaldi V, Carrillo-de-Santa-Pau E, Javierre BM, Juan D, Fraser P, Spivakov M, Valencia A, Rico D: **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity**. *Genome Biol* (1) 2016, **17**:152, https://doi.org/10.1186/s13059-016-1003-3 PMID: 27391817; PMCID: PMC4939006.

38. Newman MEJ: **Assortative mixing in networks**. *Phys Rev Lett* 2002, **89**:208701.

39. Jodkowska K, Pancaldi, *et al*.: **3D chromatin connectivity underlies replication origin efficiency in mouse embryonic stem cells**. *Nucleic Acids Res* 2022, **50**:21.

40. Richart L, *et al*.: **STAG2 loss-of-function affects short-range genomic contacts and modulates the basal-luminal**

transcriptional program of bladder cancer cells. *Nucleic Acids Res* 2021, **49**:11005-11021.

41. Ben Zouari Y, Molitor AM, Sikorska N, *et al*.: **ChiCMaxima: a robust and simple pipeline for detection and visualization of chromatin looping in Capture Hi-C**. *Genome Biol* 2019, **20**:102, https://doi.org/10.1186/s13059-019-1706-3

42. Belcastro V, *et al*.: **Transcriptional gene network inference from a massive dataset elucidates transcriptome organization and gene function**. *Nucleic Acids Res* 2011, **39**:8677-8688.

43. Soler-Oliva ME, Guerrero-Martínez JA, Bachetti V, Reyes JC: **Analysis of the relationship between coexpression domains and chromatin 3D organization**. *PLoS Comput Biol* 2017, **13**:e1005708.

44. Sutherland H, Bickmore WA: **Transcription factories: gene expression in unions?** *Nat Rev Genet* 2009, **10**:457-466.

45. Ing-Simmons E, *et al*.: **Independence of chromatin conformation and gene regulation during Drosophila dorsoventral patterning**. *Nat Genet* 2021, **53**:487-499.

46. Ghavi-Helm Y, *et al*.: **Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression**. *Nat Genet* 2019, **51**:1272-1282.

47. Li T, *et al*.: **OCEAN-C: mapping hubs of open chromatin interactions across the genome reveals gene regulatory networks**. *Genome Biol* 2018, **19**:54.

48. Sumida, N. et al. The ultra-sensitive Nodewalk technique identifies stochastic from virtual, population-based enhancer hubs regulating MYC in 3D: Implications for the fitness of cancer cells. 286583 Preprint BioRxiv at https://doi.org/10.1101/286583 (2018).(this is a preprint).

49. Sumida N, *et al*.: **MYC as a driver of stochastic chromatin networks: implications for the fitness of cancer cells**. *Nucleic Acids Res* 2020, **48**:10867-10876.

50. Sandhu KS, *et al*.: **Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development**. *Genes Dev* 2009, **23**:2598-2603.

51. •• Karbalayghareh A, Sahin M, Leslie CS: **Chromatin interaction aware gene regulatory modeling with graph attention networks**. *Genome Res* (5) 2022, **32**:930-944, https://doi.org/10.1101/gr.275870.121 Epub PMID: 35396274; PMCID: PMC9104700.
This paper proposes a graph neural network approach to relate chromatin structure and expression.

52. Xu, D. et al. Recapitulation of patient-specific 3D chromatin conformation using machine learning and validation of identified enhancer-gene targets. bioRxiv (2021) doi:10.1101/2021.11.16.468857.(this is a pre-print).

53. Yan Y, Li Z, Li Y, Wu Z, Yang R: **Correlated evolution of large DNA fragments in the 3D genome of Arabidopsis thaliana**. *Mol Biol Evol* 2020, **37**:1621-1636.

54. Pancaldi V: **Chromatin network analyses: towards structure-function relationships in epigenomics**. *Front Bioinform* 2021, **1**:742216.

55. Tirosh I, Reikhav S, Sigal N, Assia Y, Barkai N: **Chromatin regulators as capacitors of interspecies variations in gene expression**. *Mol Syst Biol* 2010, **6**:435, https://doi.org/10.1038/msb.2010.84 PMID: 21119629; PMCID: PMC3010112.

56. Hagai T, *et al*.: **Gene expression variability across cells and species shapes innate immunity**. *Nature* 2018, **563**:197-202.

57. Sengupta K, *et al*.: **Multi-scale phase separation by explosive percolation with single-chromatin loop resolution**. *Comput Struct Biotechnol J* 2022, https://doi.org/10.1016/j.csbj.2022.06.063

58. Malod-Dognin N, Pancaldi V, Valencia A, Pržulj N: **Chromatin network markers of leukemia**. *Bioinformatics* 2020, **36**:i455-i463.

59. Brackley CA, Gilbert N, Michieletto D, *et al*.: **Complex small-world regulatory networks emerge from the 3D organisation of the human genome**. *Nat Commun* 2021, **12**:5756, https://doi.org/10.1038/s41467-021-25875-y

60. Wang H, Yang J, Zhang Y, Wang J: **Discover novel disease-associated genes based on regulatory networks of long-range chromatin interactions**. *Methods* 2021, **189**:22-33.

61. Hu B, *et al*.: **Neuronal and glial 3D chromatin architecture informs the cellular etiology of brain disorders**. *Nat Commun* 2021, **12**:3968.

62. Mangnier L, Joly-Beauparlant C, Droit A, Bilodeau S, Bureau A: **Cis-regulatory hubs: a new 3D model of complex disease genetics with an application to schizophrenia**. *Life Sci Alliance* 2022, **5**:e202101156.

63. Espeso-Gil S, *et al*.: **A chromosomal connectome for psychiatric and metabolic risk variants in adult dopaminergic neurons**. *Genome Med* 2020, **12**:19.

64. Kivelä M, *et al*.: **Multilayer networks**. *J Complex Netw* 2014, **2**:203-271.

65. • Chiliński M, Sengupta K, Plewczynski D: **From DNA human sequence to the chromatin higher order organisation and its biological meaning: using biomolecular interaction networks to understand the influence of structural variation on spatial genome organisation and its functional effect**. *Semin Cell Dev Biol* 2021, **121**:171-185, https://doi.org/10.1016/j.semcdb.2021.08.007.
This paper introduces the concept of biomolecular interaction networks that connect different levels of regulation of cellular processes, from chromatin to phenotype.

66. Halder AK, *et al*.: **Aggregated network centrality shows non-random structure of genomic and proteomic networks**. *Methods* 2020, **181–182**:5-14, https://doi.org/10.1016/j.ymeth.2019.11.006

67. Csardi G: **The igraph software package for complex network research**. *Inter Complex Syst* 2006,.

68. A Hagberg, D Schult, P Swart, Exploring Network Structure, Dynamics, and Function using NetworkX in Proceedings of the 7th Python in Science conference (SciPy 2008), G Varoquaux, T Vaught, J Millman (Eds.), pp. 11-15.

69. Shannon P, *et al*.: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. *Genome Res* 2003, **13**:2498-2504.

70. Bastian M, Heymann S, Jacomy MG: **An open source software for exploring and manipulating networks**. *Proc Int AAAI Conf Web Soc Media* 2009, **3**:361-362.

71. De Domenico Mothers: **MuxViz: a tool for multilayer analysis and visualization of networks**. *Journal of Complex Networks* (2) 2015, **3**:159-176, https://doi.org/10.1093/comnet/cnu038

72. Albert R, Barabási A-L: **Statistical mechanics of complex networks**. *Rev Mod Phys* 2002, **74**:47-97.

73. Newman MEJ: **Networks. An introduction**. Oxford University Press; 2010, https://doi.org/10.1111/j.1468-5922.2010.01872_2.x