Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Higher order assembly: folding the chromosome

Sven A Sewitz[1], Zahra Fahmi[1] and Karen Lipkow[1,2]

The linear molecules of DNA that constitute a eukaryotic genome have to be carefully organised within the nucleus to be able to correctly direct gene expression. Microscopy and chromosome capture methods have revealed a hierarchical organisation into territories, domains and subdomains that ensure the accessibility of expressed genes and eventually chromatin loops that serve to bring gene enhancers into proximity of their target promoters. A rapidly growing number of genome-wide datasets and their analyses have given detailed information into the conformation of the entire genome, allowing evolutionary insights, observations of genome rearrangements during development and the identification of new gene-to-disease associations. The field is now progressing into using computational models of genome dynamics to investigate the mechanisms that shape genome structure, placing increasing importance on the role of chromatin associated proteins for this process.

Addresses
[1] Nuclear Dynamics Programme, Babraham Institute, Cambridge CB22 3AT, UK
[2] Cambridge Systems Biology Centre, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK

Corresponding author: Lipkow, Karen (karen.lipkow@babraham.ac.uk, KL280@cam.ac.uk)

Current Opinion in Structural Biology 2017, **42**:162–168

This review comes from a themed issue on **Folding and binding**

Edited by **Jane Clarke** and **Rohit V Pappu**

For a complete overview see the Issue and the Editorial

Available online 8th March 2017

http://dx.doi.org/10.1016/j.sbi.2017.02.004

0959-440X/© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

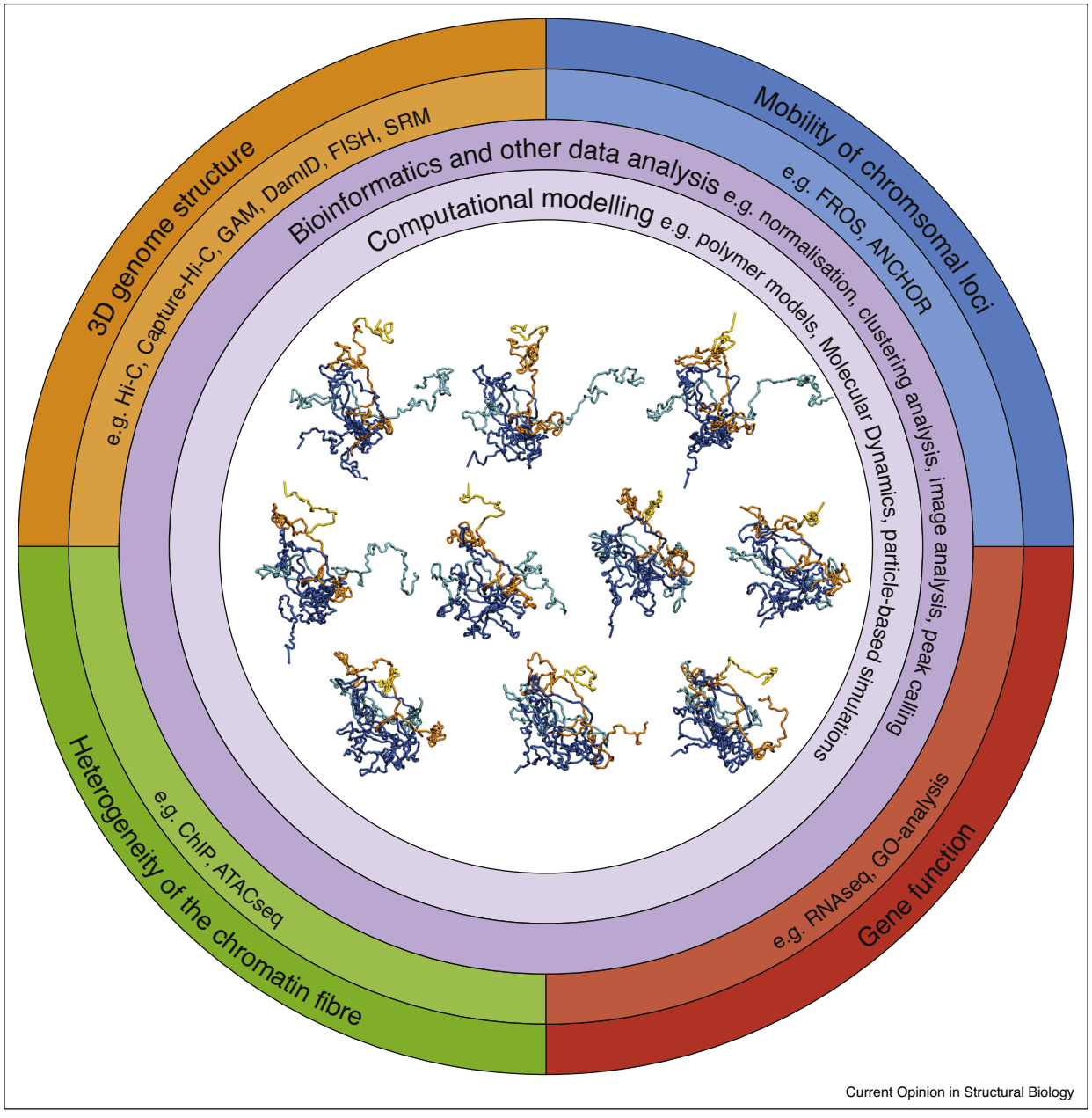
The linear DNA molecule of eukaryotic genomes has to be compacted into the nucleus by a factor of over several hundred thousand (>300 000 fold for humans), even during interphase. This is an astonishing level of compaction. At the same time, the genome remains highly organised on several scales. It is now clear that this organisation is both non-random and vital. Even subtle mistakes in arrangement can lead to severe developmental defects [1•] or disease [2,3•]. This organisation, which consists of several hierarchies, has been studied intensively by microscopy, and increasingly by conformation capture methods and computational methods (Figure 1) (see Ref. [4] for a technical review of the experimental approaches and data analysis methods). The dynamics of the genome are also being investigated at ever increasing resolution [5••]. While significant amounts of data have been collected describing the details of genome organisation, the mechanisms that shape it are only starting to emerge.

Domain structure of genomes

In the cells of higher eukaryotes, the largest structural units are chromosome territories [6]: Every interphase chromosome has its preferred location, and preferred neighbouring chromosomes, within the nucleus. Localised experimental proof came first from Fluorescence *In-situ* Hybridisation (FISH) work on human lymphocyte nuclei, which showed consistent locations of the short arms of chromosome 18 and 19 [7]. Each chromosome territory is separated by an interchromatin compartment (IC), a small uninterrupted space between chromosomes. While territories can be visualised, their boundaries are not rigid but changeable [8]. Contacts between chromosomes, or *trans* interactions, occur frequently and in certain cases coincide very well with frequencies of chromosomal translocations, linking physical contacts to observations of cytogenetic abnormalities [9]. The existence of chromosome territories on a genome-wide scale was shown comprehensively in 2009 using Hi-C [10]. This study furthermore demonstrated the spatial segregation of open, active regions from closed, inactive genomic regions: Alternating blocks of ~5 Mb come together in preferentially interacting compartments: the active A compartment and the inactive B compartment. The peripheries of territories are mostly in the A compartment [11], remain structured, and are defined by higher numbers of *trans*-chromosomal contacts [12]. At a smaller scale, topologically associated domains (TADs) are stretches of ~500 kb, which have very strong internal interactions and are insulated from other TADs [13].

The overall structure of the mammalian interphase genome is believed to be a fractal globule [14]: A chain of small globules (TADs), which organises into larger globules (regions of A/B compartments) and into territories. First derived from theoretical polymer physics [15], evidence for the fractal organisation now comes from analysis of Hi-C data [10], modelling [16,17]), and microscopy [6]. It ensures that the genome can be replicated without entanglement [17], and that each region can easily unfold and become accessible [16]. The differences in accessibility have been observed

Figure 1



The dynamics of the genome and methods of analysis.
Four main aspects of studying genome organisation are shown in the outer circle, together with select experimental techniques. To arrive at an understanding about genome conformation and dynamics, bioinformatics and other data analysis methods and increasingly computational modelling are necessary (two innermost circles). These two areas are in contact with all experimental methods. The centre shows a time series of the movement of two simulated budding yeast chromosomes within a whole-genome simulation (as in Ref. [38]). Each chromosome is tethered via the centromere (red) to the spindle pole body (not shown). The chromosome arms are colored separately, light blue (Chr IV L), dark blue (Chr IV R), yellow (Chr VIII L), and orange (Chr VIII R). The images are produced from 10 consecutive time points, with 100 simulation time steps in between any two figures and demonstrate the dynamics of chromosome conformations over a short period of time.

using superresolution microscopy (SRM), and this has also shown that parts of the genome which corresponding to different epigenetic domains are characterised by their own scaling exponent describing the packaging density of chromatin [5••].

Over the course of evolutionary time, the overall domain structure in mammals is highly conserved, while the structure within domains is more dynamic and can be modulated as a result of local sequence evolution [18•]. Also, gene paralogs are often found within the same TAD

[19], indicating that they share common regulatory mechanisms. Over the course of cellular development, it has now been shown that genome organisation undergoes stepwise changes in line with changes in chromatin modifications and gene expression [20••]. An interesting observation is that the spatial separation of euchromatin and heterochromatin correlates with the differential distribution of SINE and LINE retrotransposons [21]. The mechanisms guiding this separation are still unknown.

Specific determinants of genome organisation

While chromatin is always bound by a large number of proteins at any given moment, only a select few have been identified as being the major architectural components that determine genome organisation. At early time points during development in embryonic stem cells, the Yamanaka factor Klf4 has been shown to recruit cohesin and the mediator complex and thereby establishing TAD boundaries [22,23]. Additionally, the proteins of the Polycomb repressive complex (PRC) are essential in establishing long range interactions in early genome organisation [24–26]. The best characterised proteins involved in organising chromatin are the CCCTC-binding factor (CTCF), and the cohesin complex. CTCF has been characterised as an insulator protein, originally identified as having a repressive effect on the c-*myc* oncogene [27]. It has extensive DNA binding capability via its 11 zinc finger domains, and recognises a directional consensus sequence that has a high occurrence in mammalian cells, with over 13 000 binding sites in the human genome [28]. It functions to stabilise DNA loops jointly with cohesin, thus enabling long-range DNA–DNA interactions and creating not only large chromatin domains such as TADs, but also domains within these domains [29–32]. As another example, the Hox gene cluster can be observed to become sequentially more accessible during development, with the opening of the domain structure coinciding with enhanced gene expression [33•]. The CTCF protein is bound to several sites adjacent to the Hox gene cluster [34] and is an important determinant of gene silencing [35]. An open question in this context is how the specificity of these unique interactions is achieved, given that neither cohesin nor CTCF are able to accurately distinguish between the high number of available binding sites. This would assume a significant level of pre-organisation, for example mediated by transcription [36,37] and other mechanisms [38]. A newly proposed mechanism is the 'loop extrusion model', in which an as yet unidentified motor protein extrudes loops of DNA until it encounters two convergently oriented CTCF binding sites [39,40,41•].

The functions of gene-looping

Long range interactions between promoters and regulatory elements [26] are key to understanding the functions of gene looping. Several studies have now presented data showing that looping is associated with enhancer contacts. In both mouse and human lymphoblastoid cells, promoter interacting fragments show an enrichment of enhancer annotations [42]. Also, cell-type-specific loops were seen to coincide with cell-type-specific gene expression. While earlier studies using standard Hi-C suffered from high background interactions that masked many promoter-enhancer loops, enhanced chromosome capture methods, such as promoter-capture Hi-C [43] have shown that chromosome contacts between promoters and promoter interacting regions (PIRs) are significantly enriched for enhancer sequences. This is evident from the chromatin signature of the captured PIRs: they show enrichments for H3K4me1 and H3K27ac, classical marks that characterise enhancer regions [44]. In a study covering 17 human blood cell types, the captured interactions cluster according to cell type and offer an overall match between cell-type-specific interactions and cell-type-specific gene expression [20••]. Across all genes, the number of active enhancers that interact with a given promoter also shows a positive correlation with the level of gene expression. Using eQTL data, more than 1000 sequence variants in two cell types showed physical interactions with their target genes, highlighting the regulatory role for the detected PIRs. Extending this analysis to disease variants through Genome Wide Association Studies (GWAS), in many of which the target gene has not been identified so far, capture Hi-C had the power to newly assign more than 1800 genes to existing disease pathways [20••]. This is an indication of the significant wealth of information still to be uncovered by investigating genome organisation.

Interaction with the nuclear lamina

Apart from intra- and inter-chromosomal contacts, the chromatin fiber makes specific contacts with the nuclear lamina. These Lamin Associated Domains (LADs) are between 100 kb and 10 Mb in length, have low levels of gene expression and chromatin signatures of heterochromatin [45,46]. LADs show high interaction with Lamin B1 and are late replicating, compared to domains with low Lamin B1 interaction scores [47]. The interaction of chromatin with the lamina is not static, as single cell studies have demonstrated the dynamics of these interactions during interphase. While the mobility of individual LADs could be clearly measured, it was also constrained to regions close to the envelope, with very rare movements into neighboring euchromatin [48]. Their importance can be seen in cells with a mutated copy of Lamin A, which is a developmentally regulated constituent of the inner nuclear membrane and co-localises with Lamin B1 [49]. Individuals with a *de novo* point mutation at position 1824 of Lamin A develop the severe Hutchinson-Gilford progeria (premature ageing) syndrome. The cells are characterised by a near complete loss of the characteristic A/B compartments in late passage cells,

and a concomitant loss of H3K27me3 marks at repressed genes [2•], as well as defects in DNA double strand repair [50]. The severity and broad range of disease phenotypes, which include nuclear shape abnormalities, changes in gene expression and epigenetic regulation, telomere shortening, and genome instability, highlight the complex and wide ranging connections between genome organisation and nuclear functions.

Hi-C analysis and reconstruction of 3D genome structures

In Hi-C, as in all high-throughput and next-generation sequencing techniques, carefully conducted experiments are only the first part of the process, with the subsequent bioinformatic data processing and analysis often the more time-intensive part. The sequencing reads need to be mapped to the genome, quality checked, experimental artifacts removed and experimental biases controlled for [51]. This filtered data then forms the basis for further analysis such as identification of significantly increased interactions or TAD calling. A multitude of software tools have been developed, which can do one or several steps [52,4]. These tools use different approaches and are based on different assumptions, and as a consequence do give differing results. At the moment, there is no accepted gold-standard suite of analysis tools. It is the responsibility of each team to judge which approach is best suited to their data and research question.

Of particular appeal is the reconstruction of the 3D genome structure. The data which chromosome capture techniques produce are usually represented as contact maps, based on the interaction frequencies of the underlying DNA fragments. While these contact maps give global or local views representing chromatin conformation, they offer only limited insight into overall 3D organisation. The data presented in such a way comes only from pairwise interactions, and cannot highlight the proximity of more than two loci very well. It is thus very unintuitive to represent genomic neighbourhoods, and to indicate, for example which two TADs are in close spatial proximity or which regions are located on the periphery of territories. Strictly speaking, 3D genome reconstruction is only possible from single cell Hi-C data. Ensemble Hi-C experiments are performed on a very large population of cells, mostly upwards of 10^6. The contacts from such experiments are distributed over as many cells, and might only be present in a subset of them at a given time. Any consensus structure from ensemble Hi-C is in actual fact a probabilistic mean, and as such could be the result of averaging two or more very distinct underlying structures. Using computational reconstructions from single cell Hi-C data, it was seen that some chromosomal regions displayed large conformational differences between cells, while other features were present in most cells analysed [12]. This gives a cautionary note about being overly confident when using ensemble Hi-C data to reconstruct genome structures. At the same time, ensemble Hi-C still has its place; for once, because it allows higher depth and resolution. The need for accurate reconstruction methods is high, as having a 3D model aids understanding of genome organisation and enables the formulation of new questions. As a result, new methods are being introduced or updated on a regular basis, which reconstruct overall 3D genome structures from pairwise interaction data, optimising for either distance or probability [52,53]. An alternative approach is to embrace the diversity of structures: By modelling chromosomes as polymers and optimising the simulation parameters until the ensemble of simulated structures reproduce the same contact map as the experimentally determined one, an ensemble of structures is obtained that can be characterised (see centre of Figure 1 for examples). The models can be validated by introducing experimentally verifiable modifications [54].

Chromosome dynamics

Chromosome capture techniques, however powerful they are, capture snapshots of a genome's conformation in time. In order to fully understand the mechanisms that shape genome organisation, it is necessary to get accurate information about the dynamics of the genome. To achieve this, most studies have used microscopy to follow individually labeled gene loci, using either FROS (Fluorescent Repressor-Operator Systems) or more versatile derivatives thereof [55], to measure, for example, Mean Square Displacement (MSD) of the chromatin fiber [56,57]. As yet, there are no genome-wide experimental methods available to determine the mobilities of all gene loci at once. Our current understanding is therefore still very limited, and has put a large emphasis on computational models of genome dynamics. As with the above mentioned 3D reconstruction approaches, a number of methods have been developed to address this question. Computational models have been used to reproduce experimentally observed contact maps from the ground up, in higher eukaryotic systems [54,58,59,60•] as well as yeast [38,61–63], and to model the movement of chromatin-associated proteins [64]. In some cases this has resulted in being able to confirm assumptions that seem immediately intuitive [65], such as that placing protein binding sites at pre-determined positions leads to defined compaction. At the other end of the spectrum, computational models that reproduce experimental data to a high degree suffer from the requirement of iteratively optimising a large number of variables (up to 27), from which it is non-trivial to make the connection to physical properties of the chromatin fiber [66]. Themes that are emerging from these studies is that there is a functional connection between epigenetic marks and 3D genome organisation [58], the mechanistic details of which still remain to be determined.

Conclusion and outlook: integrating and re-evaluating our understanding of the mechanisms of chromosome folding

The field of genome architecture has made great advances within its relatively short existence. One reason is that it has benefitted from next-generation-sequencing techniques, improved microscopy methods and great strides in bioinformatics, other data analysis methods and computational modelling, sometimes performed in the same lab (Figure 1). None the less, or because of the use of diverse techniques, this has lead to discrepancies between different types of results [67,68]. These cases indicate areas where our understanding is incomplete and give us the opportunity to re-evaluate the conceptual models we use to understand how genome organisation is brought about. An area that is attracting increasing attention is to include the biophysical properties of chromatin into computational models. It is now evident that the view of chromatin as a homopolymer is a stark oversimplification. It misses the vast and diverse multitude of chromatin associated proteins that are in constant flux, binding and unbinding from DNA. The steady state concentration of these proteins is hard to determine accurately, but cannot be discounted when addressing the question of the behaviour of chromatin dynamics. As large and quantitative datasets are becoming available [69], it is upon us in the scientific community to include this knowledge into computational models, as is starting to happen [38,60•,70]. The resulting insights will significantly improve our understanding of genome dynamics, and begin to tease out the mechanisms the cell uses to control genome organisation during the many cellular processes that are intertwined with it.

Funding

This work was supported by a Royal Society University Research Fellowship (KL); Microsoft Research Faculty Fellowship (KL); the Babraham Institute (KL & SAS); the Biotechnology and Biological Sciences Research Council UK [grant number BBS/E/B/000C0405] (to Peter Fraser) and the Islamic Development Bank (ZF).

Acknowledgements

We thank Csilla Varnai for very helpful discussions. Due to the limit of this review, we regret not being able to reference many excellent contributions of our colleagues on the subject.

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. • Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R *et al.*: **Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions**. *Cell* 2015, **161**:1012-1025.

This study investigates naturally occuring structural genome variants, and shows how these disrupt genome organisation, leading to ectopic activation of developmental genes.

2. McCord RP, Nazario-Toole A, Zhang H, Chines PS, Zhan Y, Erdos MR, Collins FS, Dekker J, Cao K: **Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome**. *Genome Res* 2013, **23**:260-269.

3. • Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suvà ML, Bernstein BE: **Insulator dysfunction and oncogene activation in IDH mutant gliomas**. *Nature* 2016, **529**:110-114.

This paper shows how a mutation leading to the reduced function of a CpG demethylase results in the hypermethylation of a CTCF binding site. This leads to genome misfolding and eventually the spurious activation of an oncogene via a mislocated enhancer.

4. Schmitt AD, Hu M, Ren B: **Genome-wide mapping and analysis of chromosome architecture**. *Nat Rev Mol Cell Biol* 2016, **17**:743-755.

5. •• Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu C-T, Zhuang X: **Super-resolution imaging reveals distinct chromatin folding for different epigenetic states**. *Nature* 2016, **529**:418-422.

The authors use super-resolution microscopy and computational modelling to investigate the mobility of chromatin in *Drosophila* Kc167 cells, revealing different biophysical properties of active, Polycomb-repressed or largely unmodified chromatin.

6. Cremer T, Cremer M: **Chromosome territories**. *Cold Spring Harb Perspect Biol* 2010, **2**:a003889.

7. Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA: **Differences in the localization and morphology of chromosomes in the human nucleus**. *J Cell Biol* 1999, **145**:1119-1131.

8. Fraser P, Bickmore W: **Nuclear organization of the genome and the potential for gene regulation**. *Nature* 2007, **447**:413-417.

9. Bickmore WA, Teague P: **Influences of chromosome size, gene density and nuclear position on the frequency of constitutional translocations in the human population**. *Chromosome Res* 2002, **10**:707-715.

10. Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO *et al.*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **326**:289-293.

11. Boyle S, Rodesch MJ, Halvensleben HA, Jeddeloh JA, Bickmore WA: **Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis**. *Chromosome Res* 2011, **19**:901-909.

12. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A, Fraser P: **Single-cell Hi-C reveals cell-to-cell variability in chromosome structure**. *Nature* 2013, **502**:59-64.

13. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions**. *Nature* 2012, **485**:376-380.

14. Grosberg A, Rabin Y, Havlin S, Neer A: **Crumpled globule model of the three-dimensional structure of DNA**. *EPL* 1993, **23**:373.

15. Grosberg AY, Nechaev SK, Shakhnovich EI: **The role of topological constraints in the kinetics of collapse of macromolecules**. *J Phys France* 1988, **49**:2095-2100.

16. Mirny LA: **The fractal globule as a model of chromatin architecture in the cell**. *Chromosome Res* 2011, **19**:37-51.

17. Goldar A, Arneodo A, Audit B, Argoul F, Rappailles A, Guilbaud G, Petryk N, Kahli M, Hyrien O: **Deciphering DNA replication dynamics in eukaryotic cell populations in relation with their averaged chromatin conformations**. *Sci Rep* 2016, **6**:22469.

18. • Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S: **Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture**. *Cell Rep* 2015, **10**:1297-1309.

An evolutionary analysis of the domain architecture in four mammalian species is conducted, showing both high levels of conservation for a large subset of domains, and identifying divergent domains, demarkated by either conserved or variable CTCF binding.

19. Ibn-Salem J, Muro EM, Andrade-Navarro MA: **Co-regulation of paralog genes in the three-dimensional chromatin architecture**. *Nucleic Acids Res* 2017, **45**:81-91.

20. •• Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, Cairns J, Wingett SW, Várnai C, Thiecke MJ *et al.*: **Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters**. *Cell* 2016, **167**:1369-1384 e19.
This work uses promoter capture Hi-C on 17 primary human blood cell types to link disease variants to previously unassociated genes, enabling the future identification of a large number of new causal target genes.

21. Solovei I, Thanisch K, Feodorova Y: **How to rule the nucleus: divide et impera**. *Curr Opin Cell Biol* 2016, **40**:47-59.

22. Wei Z, Gao F, Kim S, Yang H, Lyu J, An W, Wang K, Lu W: **Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency**. *Cell Stem Cell* 2013, **13**:36-47.

23. Cuylen S, Haering CH: **A new cohesive team to mediate DNA looping**. *Cell Stem Cell* 2010, **7**:424-426.

24. Denholtz M, Bonora G, Chronis C, Splinter E, de Laat W, Ernst J, Pellegrini M, Plath K: **Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization**. *Cell Stem Cell* 2013, **13**:602-616.

25. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK *et al.*: **Polycomb complexes repress developmental regulators in murine embryonic stem cells**. *Nature* 2006, **441**:349-353.

26. Schoenfelder S, Sugar R, Dimond A, Javierre B-M, Armstrong H, Mifsud B, Dimitrova E, Matheson L, Tavares-Cadete F, Furlan-Magaril M *et al.*: **Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome**. *Nat Genet* 2015, **47**:1179-1186.

27. Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G, Neiman PE, Collins SJ, Lobanenkov VV: **An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes**. *Mol Cell Biol* 1996, **16**:2802-2813.

28. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B: **Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome**. *Cell* 2007, **128**:1231-1245.

29. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M: **Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus**. *Nature* 2009, **460**:410-413.

30. Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M: **Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster**. *EMBO J* 2009, **28**:1234-1245.

31. Seitan VC, Merkenschlager M: **Cohesin and chromatin organisation**. *Curr Opin Genet Dev* 2012, **22**:93-100.

32. Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters J-M, Murrell A: **Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus**. *PLoS Genet* 2009, **5**:e1000739.

33. • Noordermeer D, Leleu M, Schorderet P, Joye E, Chabaud F, Duboule D: **Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci**. *eLife* 2014, **3**: e02557.
This work follows the step-wise decompaction of the Hox gene cluster along with the colinear expression of the genes and involves significant changes in long-range interactions.

34. Phillips-Cremins JE, Sauria MEG, Sanyal A, Gerasimova TI, Lajoie BR, Bell JSK, Ong C-T, Hookway TA, Guo C, Sun Y *et al.*: **Architectural protein subclasses shape 3D organization of genomes during lineage commitment**. *Cell* 2013, **153**:1281-1295.

35. Ferraiuolo MA, Rousseau M, Miyamoto C, Shenker S, Wang XQD, Nadler M, Blanchette M, Dostie J: **The three-dimensional architecture of Hox cluster silencing**. *Nucleic Acids Res* 2010, **38**:7472-7484.

36. Papantonis A, Cook PR: **Genome architecture and the role of transcription**. *Curr Opin Cell Biol* 2010, **22**:271-276.

37. Larkin JD, Cook PR, Papantonis A: **Dynamic reconfiguration of long human genes during one transcription cycle**. *Mol Cell Biol* 2012, **32**:2738-2747.

38. Sewitz SA, Fahmi Z, Aljebali L, Bancroft J, Brustolini OJB, Saad H, Goiffon I, Varnai C, Wingett S, Wong H, Javierre B-M, Schoenfelder S, Andrews S, Oliver SG, Fraser P, Bystricky K, Lipkow K: **Heterogeneous chromatin mobility derived from chromatin states is a determinant of genome organisation in *S. cerevisiae***. *bioRxiv* 2017, **106344** http://dx.doi.org/10.1101/106344.

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

40. Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J *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.

41. • Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA: **Formation of chromosomal domains by loop extrusion**. *Cell Reports* 2016, **15**:2038-2049.
A computational study based on publicly available Hi-C data that proposes a model for TAD formation. The key observation is that the orientation and occupancy of CTCF binding sites can determine the presence, size and position of topologically associated domains.

42. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping**. *Cell* 2014, **159**:1665-1680.

43. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, Wingett SW, Andrews S, Grey W, Ewels PA *et al.*: **Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C**. *Nat Genet* 2015, **47**:598-606.

44. Creyghton MP, Cheng AW: **Histone H3K27ac separates active from poised enhancers and predicts developmental state**. *Proc Natl Acad Sci U S A* 2010, **107**:21931-21936.

45. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W *et al.*: **Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions**. *Nature* 2008, **453**:948-951.

46. Meuleman W, Peric-Hupkes D, Kind J, Beaudry J-B, Pagie L, Kellis M, Reinders M, Wessels L, van Steensel B: **Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence**. *Genome Res* 2013, **23**:270-280.

47. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK *et al.*: **Topologically associating domains are stable units of replication-timing regulation**. *Nature* 2014, **515**:402-405.

48. Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B: **Single-cell dynamics of genome-nuclear lamina interactions**. *Cell* 2013, **153**:178-192.

49. Dittmer TA, Misteli T: **The lamin protein family**. *Genome Biol* 2011, **12**:222.

50. Gonzalo S, Kreienkamp R: **DNA repair defects and genome instability in Hutchinson–Gilford progeria syndrome**. *Curr Opin Cell Biol* 2015, **34**:75-83.

51. Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, Andrews S: **HiCUP: pipeline for mapping and processing Hi-C data**. *F1000Res* 2015, **4**:1310.

52. Shavit Y, Merelli I, Milanesi L, Lio' P: **How computer science can help in understanding the 3D genome architecture**. *Brief Bioinform* 2016, **17**:733-744.

53. Serra F, Baù D, Filion G, Marti-Renom MA: **Structural features of the fly chromatin colors revealed by automatic three-dimensional modeling**. *bioRxiv* 2016 http://dx.doi.org/10.1101/036764.

54. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, Dekker J, Tiana G, Heard E: **Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription**. *Cell* 2014, **157**:950-963.

55. Saad H, Gallardo F, Dalvai M, Tanguy-le-Gac N, Lane D, Bystricky K: **DNA dynamics during early double-strand break processing revealed by non-intrusive imaging of living cells**. *PLoS Genet* 2014, **10**:e1004187.

56. Hajjoul H, Mathon J, Ranchon H, Goiffon I, Mozziconacci J, Albert B, Carrivain P, Victor J-M, Gadal O, Bystricky K *et al.*: **High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome**. *Genome Res* 2013, **23**:1829-1838.

57. Bystricky K: **Chromosome dynamics and folding in eukaryotes: insights from live cell microscopy**. *FEBS Lett* 2015, **589**:3014-3022.

58. 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.

59. Kurkcuoglu O, Bates PA: **Mechanism of cohesin loading onto chromosomes: a conformational dynamics study**. *Biophysj* 2010, **99**:1212-1220.

60. • Brackley CA, Brown JM, Waithe D, Babbs C, Davies J, Hughes JR, Buckle VJ, Marenduzzo D: **Predicting the three-dimensional folding of cis-regulatory regions in mammalian genomes using bioinformatic data and polymer models**. *Genome Biol* 2016, **17**:1-16.

This work simulates the mouse globin locus computationally, and incorporates experimentally determined values for CTCF binding, histone modification and DNase hypersensitivity to guide the modelling output.

61. Wong H, Marie-Nelly H, Herbert S, Carrivain P, Blanc H, Koszul R, Fabre E, Zimmer C: **A predictive computational model of the dynamic 3D interphase yeast nucleus**. *Curr Biol* 2012, **22**:1881-1890.

62. Gong K, Tjong H, Zhou XJ, Alber F: **Comparative 3D genome structure analysis of the fission and the budding yeast**. *PLoS One* 2015, **10**:e0119672.

63. Albert B, Léger-Silvestre I, Normand C, Gadal O: **Nuclear organization and chromatin dynamics in yeast: biophysical models or biologically driven interactions?** *BBA—Gene Regul Mech* 1819, **2012**:468-481.

64. Schmidt HG, Sewitz S, Andrews SS, Lipkow K: **An integrated model of transcription factor diffusion shows the importance of intersegmental transfer and quaternary protein structure for target site finding**. *PLoS ONE* 2014, **9**:e108575.

65. Barbieri M, Fraser J, Lavitas L-M, Chotalia M, Dostie J, Pombo A, Nicodemi M: **A polymer model explains the complexity of large-scale chromatin folding**. *Nucleus* 2013, **4**:267-273.

66. Di Pierro M, Zhang B, Aiden EL, Wolynes PG, Onuchic JN: **Transferable model for chromosome architecture**. *Proc Natl Acad Sci U S A* 2016, **113**:12168-12173.

67. Dekker J: **Mapping the 3D genome: aiming for consilience**. *Nat Rev Mol Cell Biol* 2016, **17**:741-742.

68. Williamson I, Berlivet S, Eskeland R, Boyle S, Illingworth RS, Paquette D, Dostie J, Bickmore WA: **Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization**. *Genes Dev* 2014, **28**:2778-2791.

69. Venters BJ, Wachi S, Mavrich TN, Andersen BE, Jena P, Sinnamon AJ, Jain P, Rolleri NS, Jiang C, Hemeryck-Walsh C *et al.*: **A comprehensive genomic binding map of gene and chromatin regulatory proteins in *Saccharomyces***. *Mol Cell* 2011, **41**:480-492.

70. Sewitz S, Lipkow K: **Systems biology approaches for understanding genome architecture**. In *Methods in Molecular Biology*, 1431. Edited by Leake MC. Chromosome Architecture: Methods and Protocols; 2016:109-126.