



HAL
open science

The 3D genome: From fundamental principles to disease and cancer

David Umlauf, Raphaël Mourad

► **To cite this version:**

David Umlauf, Raphaël Mourad. The 3D genome: From fundamental principles to disease and cancer. *Seminars in Cell and Developmental Biology*, 2019, 90, pp.128 - 137. 10.1016/j.semcdb.2018.07.002 . hal-03485092

HAL Id: hal-03485092

<https://hal.science/hal-03485092v1>

Submitted on 20 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

The 3D Genome: From fundamental principles to Disease and Cancer

David Umlauf^{1,*} and Raphaël Mourad^{2,*}

May 14, 2018

¹ LBME, Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, UPS, 31062 Toulouse, France

² LBCMCP, Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, UPS, 31062 Toulouse, France

* Correspondence: umlauf@ibcg.biotoul.fr; raphael.mourad@ibcg.biotoul.fr.

The authors equally contributed.

Key words: 3D genome; Epigenetics; Hi-C; GWAS; Cancer; Chromosomal rearrangement.

Abstract

In higher eukaryotes, the three-dimensional (3D) organization of the genome is intimately related to numerous key biological functions including gene expression, DNA repair and DNA replication regulations. Alteration of this 3D organization is detrimental to the organism and can give rise to a broad range of diseases such as cancers. Here, we review recent advances in the field. We first describe how the genome is packed in 3D to form chromosome territories, compartments and domains in a hierarchical manner. We also give an overview of the recent techniques that allowed to map the genome in 3D up to the kilobase resolution. We then discuss potential mechanisms by which genome misfolding can affect proper gene expression by distal enhancers, or can impact DNA repair leading to chromosomal rearrangements.

1 Introduction

In higher eukaryotes, development or tissue-specificity is achieved through the establishment and maintenance of highly regulated gene expression programs. Perturbations of these programs is detrimental to the organisms and may lead to disease. They can be the consequence of genetic alterations such as chromosomal translocations, deletions, insertions, mutations or DNA damages but also involve modifications of the epigenomes. In all cases, there exists an interplay between these deregulations and genome architecture. The nucleus of interphasic cells is complex and highly compartmentalized. Understanding how the genome is packed into its three dimensional space to allow timely gene expression is a question of outstanding importance and, as a matter of fact, is an intense research field. Decades of imaging studies have been instrumental in deciphering basic principles of nuclear architecture. Additionally, techniques such as chromosome conformation capture coupled to high throughput sequencing (Hi-C) have unveiled the different levels of genome organization to an extent that was probably not anticipated. Our understanding of how genes are positioned with respect to their regulatory elements within an adequate chromatin environment to achieve a correct transcriptional output is now starting to emerge. In this review, we will briefly describe fundamental concepts of nuclear architecture before stepping into the

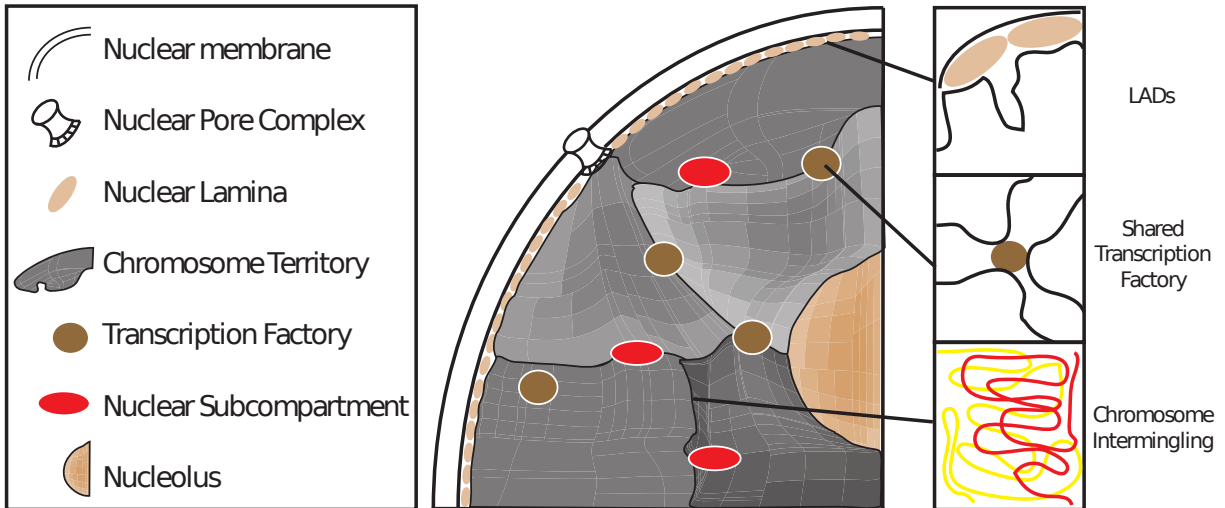


Figure 1. Schematic portion of a nucleus.

more recent data obtained by Hi-C. Finally, we will overview studies revisiting classic genetic alterations and pathologies in light of these new findings.

2 Nuclear architecture in higher eukaryotes

2.1 Chromosome territories

In higher eukaryotes, chromosomes are not randomly distributed in the nuclear space but instead occupy discrete volumes called chromosome territories (CTs) that can be considered as the highest level of chromatin organization [1] (Figure 1). CTs have a radial distribution according to their gene density. Gene-rich chromosomes tend to occupy a more central position in the nucleus, whereas gene-poor chromosomes are localized preferentially at the nuclear periphery. This trend seems to be evolutionarily conserved (see for example [2]). Within the nucleus, CTs are also organized with respect to the nuclear lamina and the perinucleolar surface through specific genomic loci called lamina-associated domains (LADs) and nucleolar-associated domains (NADs), respectively. These regions represent about 45% of the genome and are enriched with condensed chromatin, repeated elements or inactive genes [3]. Interestingly, similarly to heterochromatin, LADs can be classified as constitutive or tissue specific. While the former are always found at the lamina, the latter show a dynamic behaviour during cell differentiation [4]. This points to an active and functional role of LADs and NADs in organizing the genome.

2.2 Chromosome territories and nuclear subcompartments

Additionally, CT organization is intimately coupled to all nuclear processes such as transcription, replication, RNA processing and DNA repair. These functions are confined to membrane-free nuclear subcompartments visible by microscopy as punctuated foci in the nucleus (see for review [5]). For instance, transcription factories are enriched with active RNA polymerases II that are shared between genes located on the same or on different chromosomes [6]. These active genes tend to localize on the edge or outside of their respective CTs [6, 7]. Consistently, a transcription dependent chromosome intermingling has been observed in mammalian cells which argues in favour of genes from different chromosomes occupying the same transcription factories [8]. The biological significance of such observation has been extended re-

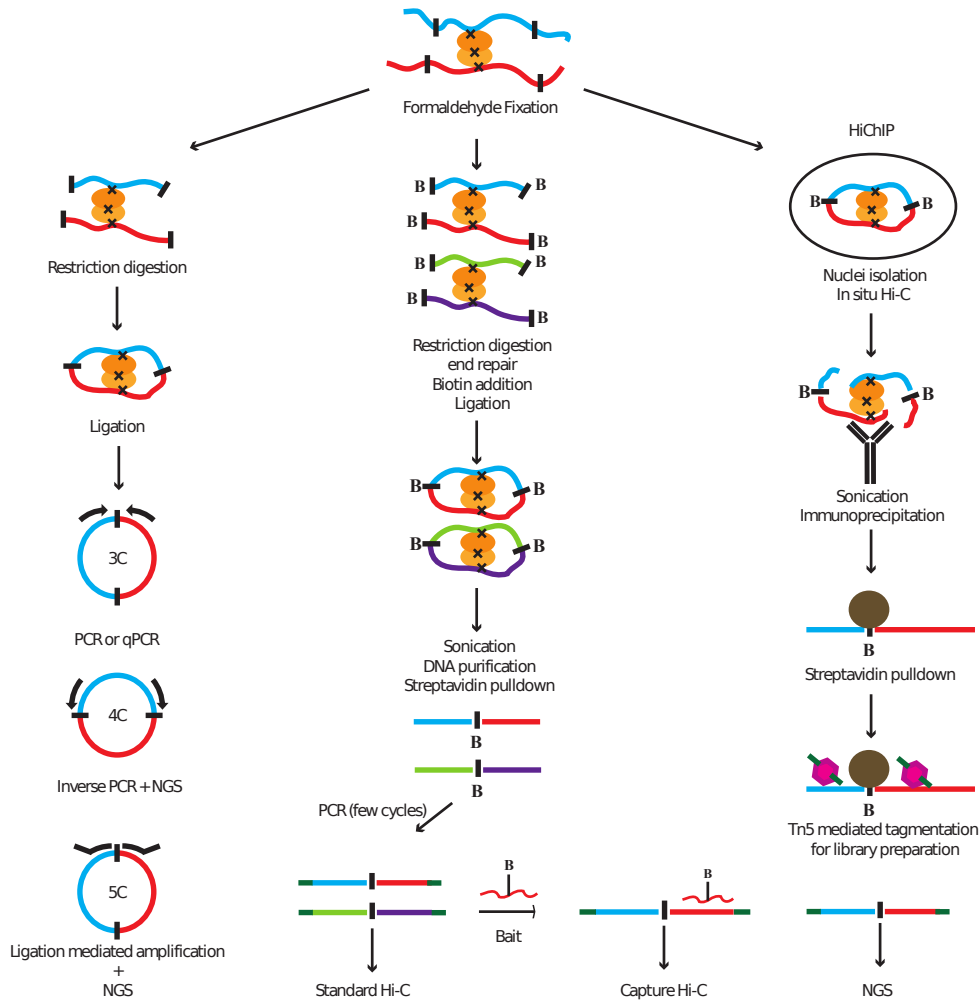


Figure 2. Illustration of chromosome conformation capture (3C) and derivatives

cently with the characterization of inactive clusters of intermingling chromosomes using both microscopy and Hi-C [9]. Taken together, the above cited studies underline the extent to which the nuclear space is highly compartmentalized. Yet, the intimate organization of the interphasic chromosomes was still lacking until the emergence of chromosome conformation capture-based techniques coupled with high throughput sequencing.

3 Chromosome conformation capture based techniques

3.1 Chromosome conformation capture (3C) and derivatives

3C is a proximity ligation-based assay that measures the relative frequency of interaction (hence the spatial proximity) between any two genomic loci within a population of cell [10]. These interactions can be specific and direct, indirect or result from random collisions of the chromatin fiber. Therefore, any 3C-based technique warrants careful controls and normalization before drawing any conclusions (see for review ref [11]). 3C has proved invaluable in formally demonstrating that specific genes and their

regulatory elements engage in looping interactions during gene expression or repression (see for example [12]). Over the years, multiple variants have been implemented (Figure 2). Among these, circularized chromosome conformation capture (4C) addresses genome-wide interactions of specific genomic baits [13,14] and has been very useful to study tissue-specific or developmentally regulated genomic interactions [15]. Chromosome conformation capture carbon copy (5C) in the other hand allows one to interrogate virtually all interactions within and/or between any genomic regions of interest through the use of primers that allow ligation mediated amplification [16]. High throughput 3C (Hi-C) measures the frequency of interactions on a genome wide scale in an unbiased manner [17]. Since the complexity of a standard Hi-C library may mask interactions of interest like those between promoters and their distal regulatory elements, additional variants aiming at selecting a subset of interactions have been designed (Figure 2). For example, capture Hi-C (CHi-C) makes use of biotinylated RNA baits to capture specific fragments. Some variants introduce an immunoprecipitation step that will select all interactions mediated by a given protein [18–20].

3.2 Hi-C is a technique of choice to study genome organization

The final output of a standard Hi-C experiment is a series of normalized contact matrices representing all interactions occurring within the same or between any two pairs of chromosomes at a given resolution (Figure 3). As expected, analysis of these matrices shows that the vast majority of the interactions occur intrachromosomally but specific interchromosomal contacts can also be detected as for example between small gene-rich chromosomes [17]. The probability $P(s)$ of contact between any two loci on the same chromosome given a distance s shows power-law decay confirming the polymeric nature of chromatin and is always greater than contacts between chromosomes. In the megabase range, $P(s)$ is compatible with a fractal globule model that could explain nicely why chromatin is able to unfold and refold easily during gene expression or repression. However, $P(s)$ displays other regimes depending on the range of s and this reflects the different levels of genome folding (see for review [21,22]). Hi-C data have to be normalized to remove intrinsic biases such as mappability, restriction fragment length or GC content near the ligated fragment ends [23,24]. Although different normalization algorithms applied to the same dataset can give rise to subtle differences in the interpretation of the results, all Hi-C data so far are in agreement with the above cited imaging studies regarding the overall organization of chromosomes within the nuclear interior. Hi-C experiments have also been successfully performed on single cells. They confirm Hi-C done using a population of cells and extend them by characterizing which level of organization present cell to cell variability or is, on the contrary, extremely stable in all cells [25]. Hi-C is therefore a powerful technique that approaches nuclear architecture in an unbiased fashion. With increasing sequencing depth, it allows one to dive deeper into the next levels of organization after chromosome territories and can virtually identify all functional interactions occurring within the nucleus of any given cell type.

4 Levels of genome organization unveiled by Hi-C

4.1 A and B compartments

In all Hi-C experiments, analysis of long-range intrachromosomal and interchromosomal contacts shows that the genome can be decomposed in two main compartments named A and B. They globally correspond to active and inactive chromatin respectively [17,23,24,26–29]. Compartments A and B tend to interact more frequently with compartments of the same type whether on the same or on different chromosomes. Compartment A is associated with actively transcribed genes, active RNA polymerase II (RNAPII), DNaseI hypersensitivity, high GC content and active histone marks. Compartment B possesses the converse characteristics and seems to be more compact on average. Interchromosomal interactions seem to be preferentially, but not exclusively, mediated by the A type compartments which is in agreement

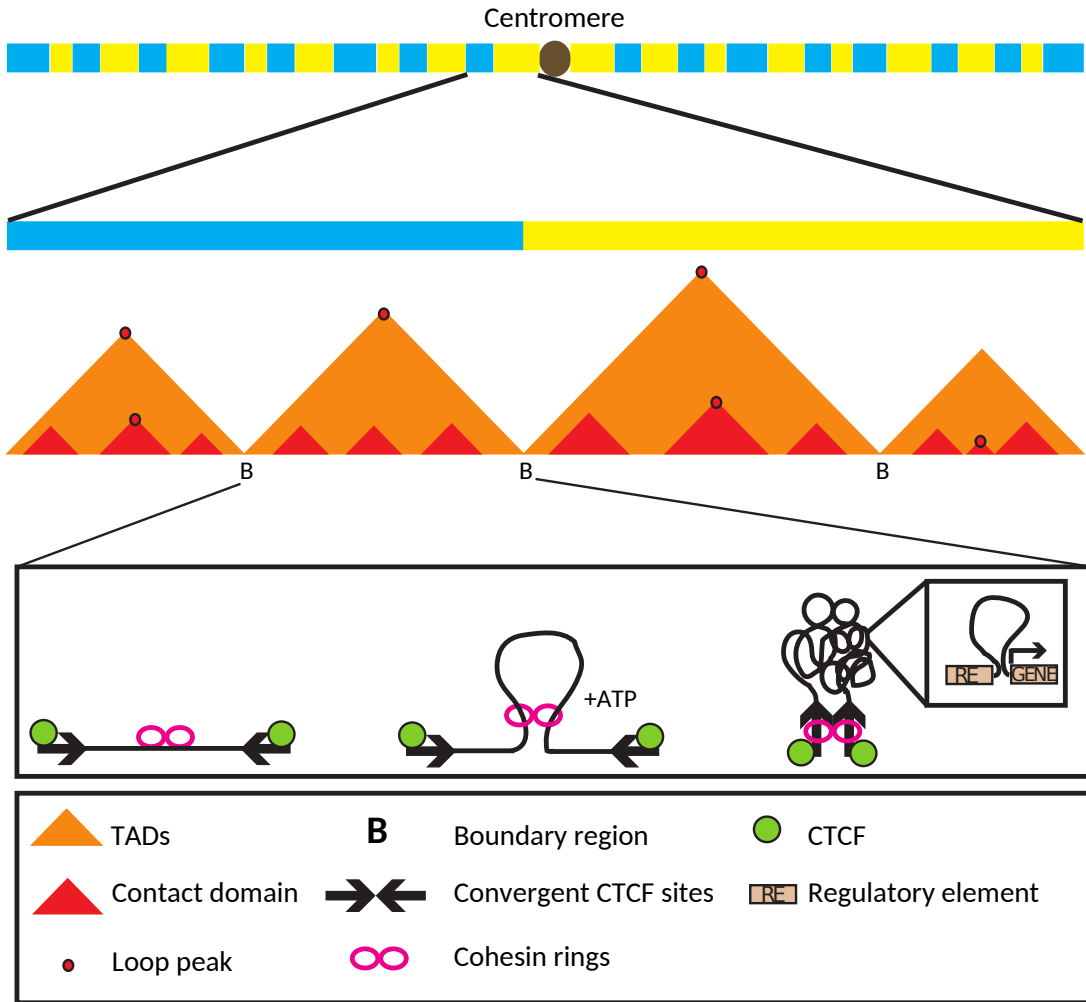


Figure 3. 3D organization of the genome in A/B compartments, topologically associating domains (TADs), contact domains and loops.

with chromosome intermingling and active genes sharing transcription factories [6, 8, 24, 28]. Additional studies have shown that it is possible to split A and B compartments in subcompartments. On the basis of their long range interactions, it was initially demonstrated that B type compartments contain centromere proximal and centromere distal clusters [24]. A high resolution Hi-C experiment [29] has since classified the A and B compartments according to distinct chromatin marks, gene activity and replication timing (Table 1). Remarkably, the distribution of A and B compartments is cell type specific and can be superimposed with early and late replication profiles respectively [17, 26, 27, 30]. Additionally, during cell differentiation, a high proportion of the genome shows A/B compartment switching [26]. Accordingly, regions with more lineage specific genes are often classified as A regions and the average expression profile of genes switching compartments is as expected: It is increased for genes switching from B to A and decreased for genes changing from A to B compartment. However, the corresponding change is mild since it only affects a subset of the genes within the compartment [26]. This suggests that compartment switching is necessary but not sufficient during cell differentiation. Consistently, in *drosophila*, a recent ultra-deep Hi-C experiment has shown that decreased interactions between A-type

Compartment	Subdivision	Features
A	A1 & A2	Gene dense, expressed genes, enriched in H3K36me3, H3K79me2, H3K27ac, H3K4me1
	A1	Replication ends beginning of S phase
	A2	Replicates into the middle of S phase, enriched in H3K9me3, lower GC content, longer genes
B	B1	Correlated to H3K27me3, depleted in H3K36me3, replication peaks during middle of S phase
	B2	Pericentromeric heterochromatin, enriched at LADs and NADs
	B3	LADs enriched, NADs depleted
	B2 & B3	Replicate during the end of S phase
	B4	Regions enriched in KRAB-ZNF superfamily genes

Table 1. Compartments A and B can be classified into finer subdivisions. (Adapted from [29]).

compartments are observed with a concomitant increase in interactions between B-type compartments during cell differentiation [31].

4.2 Topologically Associating Domains

Three milestone studies in human, mouse and *drosophila* have initially shown that genomes are partitioned into domains christened topologically associating domains (TADs) [27,32,33]. TADs correspond to regions of enriched interactions that are relatively well insulated from the rest of the genome and are sharply demarcated by boundary regions (also called borders) where a drop in local interactions is seen. In mammals, the size of TADs ranges between a hundred kilobases to a couple of megabases, whereas it is comprised between tens of kilobases to hundreds of kilobases in *drosophila* [27,33]. TADs can be aligned with epigenetics marks but do not depend on them. Indeed, on the X chromosome in mouse embryonic stem cells (mESC), TADs located within a 4.5 Mb region containing the X inactivation center align with H3K27me3 and H3K9me2. However, they remain globally intact in mESC cells lacking Eed or G9a the two histone methyl transferases that establish these marks respectively [32]. In human and mouse, most TADs are stable across cell types and, similarly to compartments A and B, are conserved between syntenic regions. However, many of the interactions that take place within TADs differ from one cell type to the other. During cell differentiation, a significant number of TADs undergo a remodelling of their intra-domain interactions [26]. For a given cell type, TADs showing an increase in interactions are more associated with active chromatin marks and TADs with decreased interactions correlate with repressive marks. In *drosophila*, the same stability of TADs is observed. However, during cell differentiation the total number of TADs is decreasing while their size increases [31].

4.3 Boundary regions play a key role in organizing TADS

In human and mice, boundary regions that delineate TADs are enriched with CTCF, cohesin, housekeeping genes, SINE elements but also tRNA genes [27]. Similarly to TADs, they are stable and conserved yet there exists instances where their exact position is subject to variation. For example, the mouse *HoxD* gene cluster is located in between two TADs and hosts a border. Both TADs contain regulatory elements that drive the expression of specific sets of *HoxD* genes during limb development. Strikingly, the boundary region is dynamic and its position accounts for the observed transcriptional profile of *HoxD* genes in proximal and distal limb bud cells [34]. Boundaries drive TADs insulation therefore facilitating intra-TADs specific interactions. Accordingly, when a boundary is genetically deleted, adjacent TADs tend to fusion and new abnormal interactions between regulatory elements and genes are created [34,35]. Additionally, genes relocated to a TADs they do not normally belong to will have ectopic interactions with regulatory elements therefore modifying their expression profile. Both situations have been shown to be causal in some developmental pathologies [35]. In *Drosophila*, borders are enriched in insulator

proteins such as CTCF, cohesin, BEAF-32, CP190 and Chromator [33,36]. As seen above, the existence of developmentally regulated TADs implies the formation of new boundaries during cell differentiation. Very often, it involves active TSS (although transcriptional activation is not sufficient) and can be CTCF independent [31].

4.4 Contact domains and looping interactions

High resolution Hi-C experiments in human and mouse have revealed smaller domains within TADs called contact domains that are locally enriched in interactions. With a median size of 185kb, they are less insulated than TADs and exhibit correlated histone marks [29]. From one cell type to the other, if this histone modification pattern changes their long-range contacts change accordingly. Interestingly, most TADs and contact domains are formed by looping interactions detected as peaks on the contact matrices which correspond to regions interacting more frequently than loci lying in between them (e.g boundary regions in the case of TADs). Between 55 – 75% of these peaks are conserved across all human cell types analyzed and about half of them are conserved in mouse syntenic regions [29]. Some loops involve a promoter-enhancer interaction that correlates positively with gene expression. In the case of cell type specific peaks, their appearance is concomitant with changes in gene expression when a promoter overlaps the peak. A high proportion of loops are anchored at two CTCF sites which present the peculiarity of being in a convergent orientation. Most of these sites ($\approx 85\%$) are bound by the CTCF, RAD21 and SMC3 proteins, the last two belonging to the cohesin complex. The demarcation of TADs and contact domains by convergent CTCF binding sites is crucial. Their deletion or inversion is able to abolish loop formation, therefore leading to adjacent TADs fusion or creation of loops between ectopically created convergent CTCF sites [35,37]. Current hypothesis suggests that TADs and contact domains are formed by loop-extrusion. In this model, two rings of cohesin complex are extruding chromatin until two convergent CTCF sites are met [37]. Indeed, recent studies have demonstrated that the related condensin complex is an ATP-dependent molecular motor that is able to perform such loop extrusion in vivo [38,39]. Moreover, the ATPase activity of cohesin is required for loop extrusion in human but not necessary for the maintenance of compartments and loops as they remain stable after ATP depletion [40].

4.5 Compartments and TADs are functionally independent units

Recent studies have challenged the view of compartments and TADs being hierarchically organized. Indeed, single cell Hi-C performed in oocytes allows the detection of the standard organization (i.e compartments, TADs and loops). However, in the zygote, male and female pronuclei display distinct characteristics. The maternal genome in the zygote is devoid of detectable compartments while TADs and loops remain [41]. The absence of compartments could be explained by the absence of transcription from the maternal genome and TADs may be established independently of transcription by loop extrusion [42]. As compartments are still detected on the paternal genome, it remains to be determined whether they are inherited from the sperm or established faster in the zygote [41]. Additionally, several lines of evidence have shown that depleting cohesin or preventing their loading onto chromatin leads to the disappearance of TADs and loops while compartments still remain [43,44]. Interestingly, in this case, a finer compartmentalization is observed [43,44] with new B like compartments arising within former A compartments [44]. This observation has led to the hypothesis that TADs could actually accommodate genomic regions that would be classified in opposite compartments otherwise [44]. This trend is not observed in B compartments. Strikingly, the disappearance of TADs and loops lead to altered transcriptional profiles that are actually milder than one would expect [43,44]. Altogether, these results suggest that compartments and TADs are functionally independent.

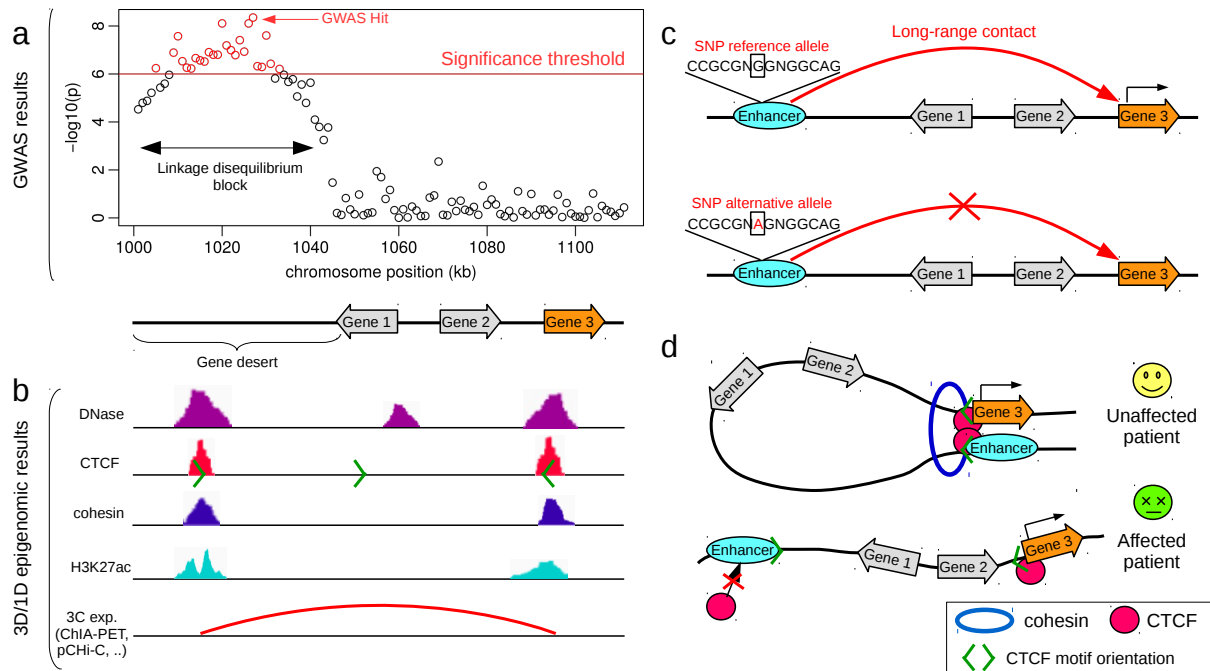


Figure 4. Post-GWAS analysis of risk loci using 3D genome information. a) GWAS mapping of single nucleotide polymorphisms (SNPs) associated with a complex disease. b) Mapping of epigenomic features including DNase, CTCF, cohesin, H3K27ac and 3C-based experiments. Green arrows represent CTCF motif orientations. c) 3C-based experiment results depending on the SNP allele. d) Functional model explaining the non-coding SNP effect.

5 Long-range contacts link SNPs to disease

Complex diseases are caused by the combined effects of multiple genes with lifestyle and environmental factors [45]. These diseases are common in the population and include heart disease, diabetes, schizophrenia and some cancers [46–49]. Over the past decade, genome-wide association studies (GWASs) have successfully identified thousands of single nucleotide polymorphisms (SNPs) associated with complex diseases in an unbiased manner [45]. However, GWASs have several limitations which hinder the precise localization of causal mutations. One limitation comes from the difficulty of gaining insight into the underlying biological mechanism, since over 95% of associated SNPs are located outside coding sequences [50]. Interestingly, more than 75% of these SNPs overlap DNase I hypersensitive sites, which suggests their association with regulatory elements [50]. Thus, a SNP might be linked to a distant gene via looping and therefore affect its expression. This would explain the discovery of thousand SNPs altering the expression of distant genes [51]. Another limitation of GWASs lies in their inability to identify the causal SNP, or the nearest SNP to the unobserved causal mutation, among a set of SNPs in strong linkage disequilibrium (LD) [52]. Because of LD, SNPs in the same haplotype block are often all significantly associated to the disease [53]. There is thus a need to connect GWAS hits to their target genes, in order to identify the causal mutations and to better understand the underlying biological mechanisms.

3C-based techniques provide a way to connect disease-associated SNPs with target genes, thus allowing a better understanding of GWAS results. Among the large number of 3C-based techniques, the most appropriate ones are those that target candidate regions of the genome rather than providing a global

view as Hi-C does. Targeted techniques include 4C [13], capture Hi-C (CHi-C) [19], chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) [18] or Hi-C chromatin immunoprecipitation (HiChIP) [20]. For instance, CHi-C analysis of CD34+ hematopoietic progenitor cells showed that interacting loci are enriched for SNPs associated with hematological parameters, diabetes and autoimmune diseases. This suggests that distal mutations may disrupt the regulation of relevant genes [19]. CTCF ChIA-PET analysis further revealed that a SNP located in the CTCF motif (CTCF-SNP) can alter the binding of CTCF and therefore reduce long-range interactions with the target gene [54]. Moreover, among the CTCF-SNP, several were found in strong linkage disequilibrium with known GWAS hits. In addition, H3K27ac HiChIP demonstrated that enhancers can fine map functional causal variants within haplotype blocks [55].

3D genome analyses can elucidate the roles of specific GWAS hits that were difficult to decipher otherwise (Figure 4). This approach can be divided into four main steps. First, we consider GWAS hits within non-coding regions, *e.g.* intergenic regions or introns. For a GWAS hit, all SNPs within the same LD block are considered as potentially causal mutations. Second, 1D epigenomic experiments (DNase-seq and ChIP-seq) and 3D genome experiments (ChIA-PET, promoter CHi-C (pChI-C), ...) are used to identify SNPs within functional regions such as enhancers and super-enhancers, and to link the SNPs to target genes. Third, long-range contact allelic imbalance between homologous chromosomes can be tested to show the impact of a functional SNP on long-range contact. Alternatively, genome editing techniques can be used to assess the impact of a SNP on long-range contacts. Fourth, a functional model is proposed to explain the role of the candidate causal SNP(s). Following this approach, the roles of several GWAS hits were recently understood. For instance, several GWASs have reproducibly associated SNPs within introns of *FTO* with obesity [56]. While studies in mice revealed that *FTO* overexpression results in obesity, no direct connection between the associated SNPs and *FTO* expression or function could be found. Instead, 4C analysis revealed that the obesity-associated SNPs within *FTO* were in long-range contact with the homeobox gene *IRX3* that was located several megabases away and was associated with body mass regulation [57]. Moreover, GWASs have consistently linked SNP rs965513, which lies in an intergenic region, to thyroid cancer [58]. This SNP is part of a linkage disequilibrium block of 33 kb that includes four SNPs overlapping enhancers. Interestingly, the latter SNPs affect enhancer activity and transcriptional factor binding. 3C assay showed that the enhancers were involved in long-range contact with the nearby genes *FOXE1* and *PTCSC2* [59].

3C-based experiments are costly and challenging to carry out. Fortunately, one can interrogate available databases before doing such experiments. There are numerous genome browser of Hi-C and other 3C-based experimental data such as Juicebox [60] and WashU Epigenome Browser [61]. In addition, several tools were developed specifically to identify candidate gene targets of a particular SNP, including the 3DSNP database [62]. For instance, the 3DSNP database integrates genotype data with 3D genome, epigenetic and chromatin signatures, DNA motifs and eQTLs, thereby accumulating evidence for gene targeting. However, most of the 3D genome data published to date were obtained from cultured cell lines, limiting their generalizability to disease studies. To tackle this issue, Hi-C data were recently generated from 14 human primary tissues, revealing strong tissue-specificity in local chromatin interactions [63]. A specific browser, HUGIn [64], was designed to visualize the former Hi-C data, thus representing a formidable resource for studying functional roles of disease-associated variants [63].

6 Genomic rearrangements are influenced by 3D genome

Chromosomal rearrangements represent a type of chromosome abnormality that involves a gross alteration of chromosome structure [65]. Chromosomal rearrangements include translocations, inversions, insertions, duplications and deletions. They can range from thousands to millions of base pairs, and can

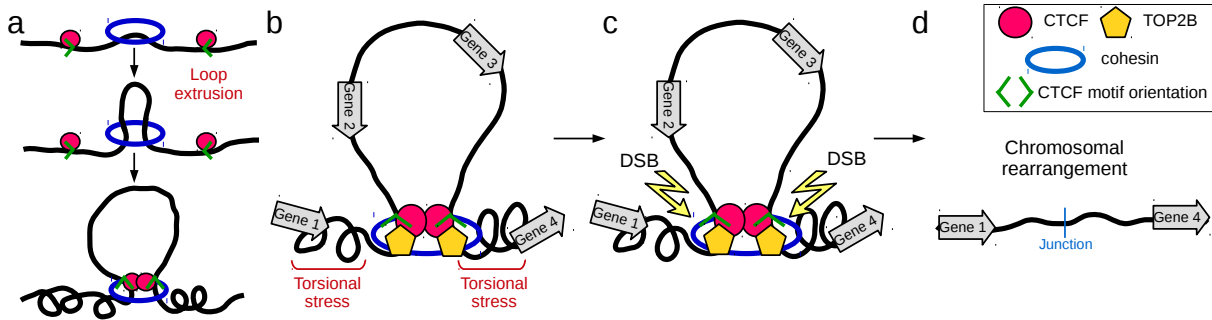


Figure 5. 3D genome dynamic induces chromosomal rearrangements. a) Loop formation by loop extrusion. Green arrows represent CTCF motif orientations. b) Torsional stress is generated at loop anchors. c) DNA topoisomerase 2 beta (TOP2B) can induce double-strand DNA breaks (DSBs), when relieving the torsional stress. d) The two broken ends can be joined together, thereby forming a chromosomal rearrangement.

cover entire gene clusters. Chromosomal rearrangements can be caused by a breakage of DNA double helices at two (or more) different locations [66]. DNA repair mechanisms in the cell then correct the DNA double-stranded breaks (DSBs) by joining broken ends back together [67]. There are two main pathways to repair DSBs: non-homologous end joining (NHEJ) and homologous recombination (HR) [68]. NHEJ directly ligates the break ends, whereas HR uses homologous sequence to guide repair. If the DNA repair is successful, the two ends of the same break are rejoined and the original DNA order is restored. But if the DNA repair fails, the two ends of different breaks are joined together, and a chromosomal rearrangement is generated. DSBs represent highly toxic lesions and can lead to cell death [69,70]. They are very frequent in cancer cells [71,72], but they are also common in normal cells. For example, high-throughput sequencing of DSBs recently mapped more than 80 thousand endogenous DSB hotspots in normal human epidermal keratinocytes (NHEK) [73,74].

An important mechanistic factor of chromosomal rearrangement is the 3D genome organization that can bring two linearly separated loci in physical proximity [75]. Supporting this hypothesis, Fudenberg *et al.* showed that the frequency of chromosomal rearrangements in cancer is related to 3D genome using Hi-C data with statistical modeling [76]. More recently, Hi-C data analysis revealed that several subfamilies of repetitive elements co-localize in 3D [77], and that SINE elements are enriched at TAD borders [27]. These results clearly suggest that nonallelic homologous recombination (NAHR) between similar sequences may be guided toward regions in physical proximity. NAHR is a form of homologous recombination that occurs between two highly similar sequences localized at two distinct loci, and which is at the origin of recurrent chromosomal rearrangements [78]. Highly similar sequences include low copy repeats (LCRs), transposons [79] or human endogenous retroviruses (HERVs) [80]. When NAHR occurs between directly oriented repeats, it generates deletions and duplications. When NAHR involves inverted repeats, it leads to the inversion of the segment in-between the repeats. It has been shown that NAHR crossover frequency in meiosis is positively associated with the flanking low-copy repeat length and inversely influenced by the distance between repeats, possibly reflecting the observed negative correlation between Hi-C contact and distance [81].

DSBs can cluster together to form repair foci that concentrate repair factors [82]. In particular, DSB clustering mostly occurs at damaged active genes during G1 [83]. Recent DSB mapping combined with Hi-C experiments revealed that DSBs often occur at loop anchors where CTCF and cohesin bind [84]. CTCF and cohesin seem to have different roles in DSB repair. On the one hand, CTCF facilitates DNA

double-strand break repair by enhancing homologous recombination repair [85]. On the other hand, cohesin can protect genes by blocking the spreading of the histone mark γ -H2AX induced by DSBs [86]. Interestingly, topoisomerase 2B (TOP2B), an enzyme known to mediate DSBs, physically interact with CTCF and cohesin at TAD borders [87]. Moreover, TOP2B is enriched at TSSs, and more generally, at open chromatin sites [88]. TOP2B is an enzyme that controls and alters the topological states of DNA. In particular, TOP2B catalyzes the transient breaking and rejoining of two strands of duplex DNA which allows the strands to pass through one another, and thus the relief of torsional stress during transcription [88]. In Figure 5, we illustrate how the 3D genome can influence the formation of chromosomal rearrangements. Loops are formed by loop extrusion, when chromatin is fed through the extrusion complex and loops are enlarged [37]. This process may generate a torsional stress at the loop anchors just outside the loop. TOP2B can then relieve the torsional stress. In rare situations, TOP2B fails and generates DSBs. If DNA repair fails, then it can lead to the formation of chromosomal rearrangements.

Because a major factor that influences the frequency of translocations between two distant loci is the frequency of their long-range contacts [75], some translocations are more frequent than others and the frequency of translocation is distance-dependent in a similar manner as Hi-C contacts, i.e. following a power-law decay [76]. The 3D genome organization plays thus a major role in generating large alteration of the chromosome structure, and consequently, in shaping the mutational landscape of tumours [89]. Under this assumption, we speculate that TAD borders are major barriers to chromosomal rearrangements, and should protect the genome in normal cells from highly deleterious mutations. Chromosomal rearrangements overlapping TAD borders should be rare in normal cells, and are likely to represent mutations with a high impact. For instance, the loss of a TAD border can induce *de novo* long-range interactions between distal limb enhancers and ectopic target genes, and cause human limb malformations [35]. In addition, silent proto-oncogenes are protected by CTCF mediated loops from active enhancers located outside. But when a loop anchor is lost, then a proto-oncogene can be activated due to inappropriate enhancer-promoter interactions [90].

7 Conclusions

The 3D organization of chromosomes is essential to numerous key processes including the regulation of gene expression by distal enhancers [91], replication-timing program [92] and DNA repair [84]. Thus, if the 3D genome is perturbed, it can affect important cellular processes and eventually lead to disease.

Specific mutations at genes encoding members of the cohesin complex and its accessory factors give rise to diseases known as cohesinopathies, characterized by craniofacial deformities, limb defects, and mental retardation [93]. Interestingly, mutations of the cohesin complex were also associated with cancer [94]. For instance, recent sequencing of acute myeloid leukemias (AMLs) revealed frequent mutations in genes encoding members of the cohesin complex [95]. In addition, mutations at nuclear lamin genes can affect structural properties of the nuclear lamina which is composed of intermediate filaments and membrane associated proteins. In particular, large chromatin domains make contact with the nuclear lamina. Such lamina-associated domains (LADs) contribute to the organization chromosomes inside the nucleus and are linked with gene repression [96]. Mutations of lamin genes cause numerous diseases called laminopathies that comprise two types of diseases: dystrophies affecting cardiac and skeletal muscles, and partial lipodystrophies [97]. However, the molecular mechanisms by which these mutations can lead to diseases have yet to be elucidated.

Mutations at functional elements such as enhancers, insulators or even TAD borders can also alter the 3D genome organization at specific loci. Thus, comparing the 3D genome between affected and unaffected patients or for matched normal-tumor pairs would allow the identification of specific long-range

contacts and TAD borders associated with genetic diseases and cancers. Supporting this experiment, the analysis of 200 human samples of colorectal cancer revealed the higher frequency of somatic mutations at CTCF-binding sites than flanking regions [98]. Moreover, the analysis of families with rare limb malformations identified chromosomal rearrangements having an impact on TAD border, and resulting in *in novo* enhancer-promoter interactions and misexpression [35]. Regarding common complex diseases, until now, no 3D genome GWAS has been carried out, which is due to the high cost of Hi-C experiments. However, since the cost of sequencing is continuously decreasing and new techniques such as capture Hi-C and HiChIP were developed to target specific loci, GWASs will become soon affordable and will help us decipher the role of non-coding variants involved in disease etiology.

Acknowledgements

This work was supported by the University Paul Sabatier and the CNRS.

References

- [1] Thomas Cremer and Marion Cremer. Chromosome territories. *Cold Spring Harbor Perspectives in Biology*, 2(3):a003889, March 2010.
- [2] Hideyuki Tanabe, Stefan Müller, Michaela Neusser, et al. Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proceedings of the National Academy of Sciences*, 99(7):4424–4429, April 2002.
- [3] Lars Guelen, Ludo Pagie, Emilie Brasset, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, 453:948–951, June 2008.
- [4] Wouter Meuleman, Daan Peric-Hupkes, Jop Kind, et al. Constitutive nuclear laminagenome interactions are highly conserved and associated with A/T-rich sequence. *Genome Research*, 23(2):270–280, February 2013.
- [5] Miroslav Dundr. Nuclear bodies: multifunctional companions of the genome. *Current Opinion in Cell Biology*, 24(3):415 – 422, June 2012. Nucleus and gene expression.
- [6] Cameron S. Osborne, Lyubomira Chakalova, Karen E. Brown, et al. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature Genetics*, 36:1065–1071, September 2004. Article.
- [7] E.V. Volpi, E. Chevret, T. Jones, et al. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *Journal of Cell Science*, 113(9):1565–1576, May 2000.
- [8] Miguel R Branco and Ana Pombo. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLOS Biology*, 4(5):e138, April 2006.
- [9] Anastasiya Belyaeva, Saradha Venkatachalapathy, Mallika Nagarajan, G. V. Shivashankar, and Caroline Uhler. Network analysis identifies chromosome intermingling regions as regulatory hotspots for transcription. *Proceedings of the National Academy of Sciences*, 114(52):13714–13719, December 2017.
- [10] Job Dekker, Karsten Rippe, Martijn Dekker, and Nancy Kleckner. Capturing chromosome conformation. *Science*, 295(5558):1306–1311, February 2002.

- [11] Job Dekker, Marc A. Marti-Renom, and Leonid A. Mirny. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Reviews Genetics*, 14(6):390–403, June 2013.
- [12] Bas Tolhuis, Robert Jan J. Palstra, Erik Splinter, Frank Grosveld, and Wouter de Laat. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Molecular Cell*, 10(6):1453–1465, December 2002.
- [13] Zhihu Zhao, Gholamreza Tavosidana, Mikael Sjölander, et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature Genetics*, 38(11):1341–1347, November 2006.
- [14] Marieke Simonis, Petra Klous, Erik Splinter, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature Genetics*, 38(11):1348–1354, November 2006.
- [15] Daan Noordermeer, Marion Leleu, Erik Splinter, et al. The dynamic architecture of *Hox* gene clusters. *Science*, 334(6053):222–225, October 2011.
- [16] Josée Dostie, Todd A. Richmond, Ramy A. Arnaout, et al. Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Research*, 16(10):1299–1309, October 2006.
- [17] Erez Lieberman-Aiden, Nynke L. van Berkum, Louise Williams, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950):289–293, October 2009.
- [18] Melissa J. Fullwood, Mei Hui H. Liu, You Fu F. Pan, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature*, 462(7269):58–64, November 2009.
- [19] Borbala Mifsud, Filipe Tavares-Cadete, Alice N. Young, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nature Genetics*, 47(6):598–606, June 2015.
- [20] Maxwell R. Mumbach, Adam J. Rubin, Ryan A. Flynn, et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nature Methods*, 13:919 EP –, Sep 2016.
- [21] Marc A. Marti-Renom and Leonid A. Mirny. Bridging the resolution gap in structural modeling of 3D genome organization. *PLOS Computational Biology*, 7(7):e1002125, July 2011.
- [22] Leonid A. Mirny. The fractal globule as a model of chromatin architecture in the cell. *Chromosome Research*, 19(1):37–51, January 2011.
- [23] Maxim Imakaev, Geoffrey Fudenberg, Rachel P. McCord, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nature Methods*, 9(10):999–1003, October 2012.
- [24] Eitan Yaffe and Amos Tanay. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nature Genetics*, 43(11):1059–1065, November 2011.
- [25] Takashi Nagano, Yaniv Lubling, Tim J. Stevens, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, 502:59–64, October 2013.
- [26] Jesse R. Dixon, Inkyung Jung, Siddarth Selvaraj, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature*, 518(7539):331–336, February 2015.

- [27] Jesse R. Dixon, Siddarth Selvaraj, Feng Yue, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376–380, May 2012.
- [28] Reza Kalthor, Harianto Tjong, Nimanthi Jayathilaka, Frank Alber, and Lin Chen. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nature Biotechnology*, 30(1):90–98, January 2012.
- [29] Suhas S. P. Rao, Miriam H. Huntley, Neva C. Durand, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7):1665–1680, December 2015.
- [30] Tyrone Ryba, Ichiro Hiratani, Junjie Lu, et al. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Research*, 20(6):761–770, June 2010.
- [31] Boyan Bonev, Netta Mendelson Cohen, Quentin Szabo, et al. Multiscale 3D genome rewiring during mouse neural development. *Cell*, 171(3):557–572, October 2017.
- [32] Elphège P. Nora, Bryan R. Lajoie, Edda G. Schulz, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 485:381–385, May 2012.
- [33] Tom Sexton, Eitan Yaffe, Ephraim Kenigsberg, et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell*, 148(3):458–472, February 2012.
- [34] Eddie Rodriguez-Carballo, Lucille Lopez-Delisle, Ye Zhan, et al. The HoxD cluster is a dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory landscapes. *Genes & Development*, 31(22):2264–2281, November 2017.
- [35] Darío G. Lupiáñez, Katerina Kraft, Verena Heinrich, et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, 161(5):1012–1025, September 2015.
- [36] Raphael Mourad and Olivier Cuvier. Computational identification of genomic features that influence 3D chromatin domain formation. *PLoS Computational Biology*, 12(5):e1004908, May 2016.
- [37] Adrian L. Sanborn, Suhas S. P. Rao, Su-Chen Huang, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences*, 112(47):E6456–E6465, November 2015.
- [38] Tsuyoshi Terakawa, Shveta Bisht, Jorine M. Eeftens, et al. The condensin complex is a mechanochemical motor that translocates along DNA. *Science*, 358(6363):672–676, November 2017.
- [39] Mahipal Ganji, Indra A. Shaltiel, Shveta Bisht, et al. Real-time imaging of DNA loop extrusion by condensin. *Science*, 360(6384):102–105, April 2018.
- [40] Laura Vian, Aleksandra Pekowska, Suhas S. P. Rao, et al. The energetics and physiological impact of cohesin extrusion. *Cell*, 173:1–14, May 2018.
- [41] Ilya M. Flyamer, Johanna Gassler, Maxim Imakaev, et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature*, 544:110114, April 2017.
- [42] Johanna Gassler, Hugo B Brandão, Maxim Imakaev, et al. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *The EMBO Journal*, 36(24):3600–3618, December 2017.
- [43] Suhas S. P. Rao, Su-Chen Huang, Brian Glenn St Hilaire, et al. Cohesin loss eliminates all loop domains. *Cell*, 171(2):305–320.e24, Oct 2017.

- [44] Wibke Schwarzer, Nezar Abdennur, Anton Goloborodko, et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, 551:51–56, November 2017.
- [45] Peter M. Visscher, Naomi R. Wray, Qian Zhang, et al. 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics*, 101(1):5–22, July 2017.
- [46] Gerald W. Dorn and Sharon Cresci. Genome-wide association studies of coronary artery disease and heart failure: where are we going? *Pharmacogenomics*, 10(2):213–223, 2009. PMID: 19207022.
- [47] Liana K. Billings and Jose C. Florez. The genetics of type 2 diabetes: what have we learned from GWAS? *Annals of the New York Academy of Sciences*, 1212(1):59–77, 2010.
- [48] Ann L. Collins and Patrick F. Sullivan. Genome-wide association studies in psychiatry: what have we learned? *The British Journal of Psychiatry*, 202(1):1–4, 2013.
- [49] Charles C. Chung, Wagner C. S. Magalhaes, Jesus Gonzalez-Bosquet, and Stephen J. Chanock. Genome-wide association studies in cancer: current and future directions. *Carcinogenesis*, 31(1):111–120, 2010.
- [50] Matthew T. Maurano, Richard Humbert, Eric Rynes, et al. Systematic localization of common disease-associated variation in regulatory dna. *Science*, 337(6099):1190–1195, 2012.
- [51] William Cookson, Liming Liang, Gonçalo Abecasis, Miriam Moffatt, and Mark Lathrop. Mapping complex disease traits with global gene expression. *Nature Reviews Genetics*, 10:184 EP –, Mar 2009. Review Article.
- [52] Sarah L. Spain and Jeffrey C. Barrett. Strategies for fine-mapping complex traits. *Human Molecular Genetics*, 24(R1):R111–R119, 2015.
- [53] Montgomery Slatkin. Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9:477 EP –, Jun 2008. Review Article.
- [54] Zhonghui Tang, Oscar J. Luo, Xingwang Li, et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell*, 163(7):1611–1627, February 2016.
- [55] Maxwell R. Mumbach, Ansuman T. Satpathy, Evan A. Boyle, et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nature Genetics*, 49:1602 EP –, Sep 2017. Article.
- [56] Christian Dina, David Meyre, Sophie Gallina, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nature Genetics*, 39(6):724–726, June 2007.
- [57] Scott Smemo, Juan J. Tena, Kyoung-Han Kim, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature*, 507(7492):371–375, March 2014.
- [58] Julius Gudmundsson, Patrick Sulem, Daniel F. Gudbjartsson, et al. Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations. *Nature Genetics*, 41:460 EP –, Feb 2009.
- [59] Huiling He, Wei Li, Sandya Liyanarachchi, et al. Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer. *Proceedings of the National Academy of Sciences*, 112(19):6128–6133, 2015.
- [60] Neva C. Durand, James T. Robinson, Muhammad S. Shamim, et al. Juicebox provides a visualization system for hi-c contact maps with unlimited zoom. *Cell Systems*, 3(1):99–101, July 2016.

- [61] Xin Zhou, Rebecca F. Lowdon, Daofeng Li, et al. Exploring long-range genome interactions using the WashU Epigenome Browser. *Nature Methods*, 10:375 EP –, Apr 2013. Correspondence.
- [62] Yiming Lu, Cheng Quan, Hebing Chen, Xiaochen Bo, and Chenggang Zhang. 3DSNP: a database for linking human noncoding SNPs to their three-dimensional interacting genes. *Nucleic Acids Research*, 45(D1):D643–D649, 2017.
- [63] Anthony D. Schmitt, Ming Hu, Inkyung Jung, et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Reports*, 17(8):2042–2059, 2018/01/03 2016.
- [64] Joshua S Martin, Zheng Xu, Alex P Reiner, et al. HUGIn: Hi-C Unifying Genomic Interrogator. *Bioinformatics*, 33(23):3793–3795, 2017.
- [65] Feng Zhang, Claudia M.B. Carvalho, and James R. Lupski. Complex human chromosomal and genomic rearrangements. *Trends in Genetics*, 25(7):298 – 307, 2009.
- [66] Torben R. Kasperek and Timothy C. Humphrey. DNA double-strand break repair pathways, chromosomal rearrangements and cancer. *Seminars in Cell & Developmental Biology*, 22(8):886 – 897, 2011. Polarized growth and movement: How to generate new shapes and structures Chromosome Recombination.
- [67] Kishore K. Chiruvella, Zhuobin Liang, and Thomas E. Wilson. Repair of double-strand breaks by end joining. *Cold Spring Harbor Perspectives in Biology*, 5(5), 2013.
- [68] Raphael Ceccaldi, Beatrice Rondinelli, and Alan D. D’Andrea. Repair pathway choices and consequences at the double-strand break. *Trends in Cell Biology*, 26(1):52–64, January 2016.
- [69] Wynand P. Roos and Bernd Kaina. DNA damage-induced cell death: From specific DNA lesions to the DNA damage response and apoptosis. *Cancer Letters*, 332(2):237 – 248, 2013. Apoptosis Targeting Drugs in Cancer.
- [70] Aaron A. Goodarzi and Penelope A. Jeggo. Chapter One - The Repair and Signaling Responses to DNA Double-Strand Breaks. In Theodore Friedmann, Jay C. Dunlap, and Stephen F. Goodwin, editors, *Advances in Genetics*, volume 82 of *Advances in Genetics*, pages 1 – 45. Academic Press, 2013.
- [71] Michael C. Haffner, Martin J. Aryee, Antoun Toubaji, et al. Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nature genetics*, 42(8):668–675, August 2010.
- [72] Peter J. O’Donovan and David M. Livingston. BRCA1 and BRCA2: breast/ovarian cancer susceptibility gene products and participants in DNA double-strand break repair. *Carcinogenesis*, 31(6):961–967, 2010.
- [73] Stefanie V. Lensing, Giovanni Marsico, Robert Hansel-Hertsch, et al. DSBCapture: in situ capture and sequencing of DNA breaks. *Nature Methods*, 13(10):855–857, October 2016.
- [74] Raphaël Mourad, Krzysztof Ginalski, Gaëlle Legube, and Olivier Cuvier. Predicting double-strand DNA breaks using epigenome marks or DNA at kilobase resolution. *Genome Biology*, 19(1):34, March 2018.
- [75] Yu Zhang, Rachel P. McCord, Yu-Jui Ho, et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell*, 148(5):908–921, March 2012.

- [76] Geoff Fudenberg, Gad Getz, Matthew Meyerson, and Leonid A. Mirny. High order chromatin architecture shapes the landscape of chromosomal alterations in cancer. *Nature Biotechnology*, 29(12):1109–1113, November 2011.
- [77] Axel Cournac, Romain Koszul, and Julien Mozziconacci. The 3D folding of metazoan genomes correlates with the association of similar repetitive elements. *Nucleic Acids Research*, 44(1):245–255, 2016.
- [78] Claudia M. B. Carvalho and James R. Lupski. Mechanisms underlying structural variant formation in genomic disorders. *Nature Reviews Genetics*, 17(4):224–238, February 2016.
- [79] Caroline Robberecht, Thierry Voet, Masoud Zamani Esteki, Beata A. Nowakowska, and Joris R. Vermeesch. Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. *Genome Research*, 23(3):411–418, 2013.
- [80] Ian M. Campbell, Tomasz Gambin, Piotr Dittwald, et al. Human endogenous retroviral elements promote genome instability via non-allelic homologous recombination. *BMC Biology*, 12(1):74, Sep 2014.
- [81] Pengfei Liu, Melanie Lacaria, Feng Zhang, et al. Frequency of nonallelic homologous recombination is correlated with length of homology: Evidence that ectopic synapsis precedes ectopic crossing-over. *The American Journal of Human Genetics*, 89(4):580 – 588, 2011.
- [82] P. Caron, J. Choudjaye, T. Clouaire, et al. Non-redundant functions of ATM and DNA-PKcs in response to DNA double-strand breaks. *Cell Reports*, 13(8):1598–1609, November 2015.
- [83] François Aymard, Marion Aguirrebengoa, Emmanuelle Guillou, et al. Genome-wide mapping of long-range contacts unveils clustering of DNA double-strand breaks at damaged active genes. *Nature Structural & Molecular Biology*, 24(4):353–361, March 2017.
- [84] Andres Canela, Yaakov Maman, Seolkyoung Jung, et al. Genome organization drives chromosome fragility. *Cell*, 170(3):507 – 521.e18, July 2017.
- [85] Khalid Hilmi, Maïka Jangal, Maud Marques, et al. CTCF facilitates DNA double-strand break repair by enhancing homologous recombination repair. *Science Advances*, 3(5):e1601898, May 2017.
- [86] Pierre Caron, François Aymard, Jason S. Iacovoni, et al. Cohesin protects genes against γ -H2AX induced by DNA double-strand breaks. *PLOS Genetics*, 8(1):e10002460, January 2012.
- [87] Liis Uusküla-Reimand, Huayun Hou, Payman Samavarchi-Tehrani, et al. Topoisomerase ii beta interacts with cohesin and ctcf at topological domain borders. *Genome Biology*, 17(1):182, Aug 2016.
- [88] Yves Pommier, Yilun Sun, Shar-yin N. Huang, and John L. Nitiss. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nature Reviews Molecular Cell Biology*, 17(11):703–721, September 2016.
- [89] Lucy R. Yates and Peter J. Campbell. Evolution of the cancer genome. *Nature Reviews Genetics*, 13(11):795–806, November 2012.
- [90] Denes Hnisz, Abraham S. Weintraub, Daniel S. Day, et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science*, 2016.
- [91] Fulai Jin, Yan Li, Jesse R. Dixon, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*, 503(7475):290–294, November 2013.

- [92] Benjamin D. Pope, Tyrone Ryba, Vishnu Dileep, et al. Topologically associating domains are stable units of replication-timing regulation. *Nature*, 515(7527):402–405, November 2014.
- [93] Rajeswari Banerji, Robert V. Skibbens, and M. Kathryn Iovine. How many roads lead to cohesinopathies? *Developmental Dynamics*, 246(11):881–888, April 2017.
- [94] Victoria K. Hill, Jung-Sik Kim, and Todd Waldman. Cohesin mutations in human cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1866(1):1 – 11, August 2016.
- [95] Swapna Thota, Aaron D. Viny, Hideki Makishima, et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood*, 124(11):1790–1798, September 2014.
- [96] Bas van Steensel and Andrew S. Belmont. Lamina-associated domains: Links with chromosome architecture, heterochromatin, and gene repression. *Cell*, 169(5):780–791, May 2017.
- [97] J. L. V. Broers, F. C. S. Ramaekers, G. Bonne, R. Ben Yaou, and C. J. Hutchison. Nuclear lamins: Laminopathies and their role in premature ageing. *Physiological Reviews*, 86(3):967–1008, 2006. PMID: 16816143.
- [98] Riku Katainen, Kashyap Dave, Esa Pitkänen, et al. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nature Genetics*, 47:818–821, June 2015.