

Methods for mapping 3D chromosome architecture

Rieke Kempfer ^{1,2*} and Ana Pombo ^{1,2*}

Abstract | Determining how chromosomes are positioned and folded within the nucleus is critical to understanding the role of chromatin topology in gene regulation. Several methods are available for studying chromosome architecture, each with different strengths and limitations. Established imaging approaches and proximity ligation-based chromosome conformation capture (3C) techniques (such as DNA-FISH and Hi-C, respectively) have revealed the existence of chromosome territories, functional nuclear landmarks (such as splicing speckles and the nuclear lamina) and topologically associating domains. Improvements to these methods and the recent development of ligation-free approaches, including GAM, SPRITE and ChIA-Drop, are now helping to uncover new aspects of 3D genome topology that confirm the nucleus to be a complex, highly organized organelle.

Chromosome territories

The nuclear volumes occupied by each specific chromosome. Chromosomes tend to interact predominantly within themselves and occupy distinct regions within the interphase nucleus.

Chromosomal compartments

Chromosomes fold into distinct subcompartments, which correlate with transcriptional activity (A compartment) or repression (B compartment). The A and B compartments are defined by Hi-C contact frequencies.

The nucleus of human cells harbours 46 densely packed chromosomes. Chromosomes are folded into hierarchical domains at different genomic scales, which likely enable efficient packaging and organize the genome into functional compartments. Chromosomes occupy distinct positions within the nucleus, called chromosome territories, which are partitioned into chromosomal compartments, and further partitioned into topologically associating domains (TADs) and chromatin loops mediated by CCCTC-binding factor (CTCF) or enhancer–promoter contacts (FIG. 1). Chromatin folding is a major feature of gene regulation and dynamically changes in development and disease^{1–4}. Transcriptional control is mediated through physical contacts between enhancers and target genes, which occur via loop formation between the respective DNA elements. Functional loops between regulatory regions and genes are thought to occur predominantly within TADs. The expression of genes can also be influenced by their positioning relative to spatial landmarks inside the nucleus that are enriched for specific biochemical activities, such as the nuclear lamina. The disruption of enhancer–gene contacts and alteration of nuclear subcompartments play important roles in disease, including congenital disorders and cancer. Importantly, many disease-associated mutations of the linear genomic sequence can only be understood by considering their 3D conformation in nuclear space.

Advances in our understanding of chromosome folding have been restricted by a lack of approaches that can map chromatin contacts genome-wide while simultaneously retrieving spatial information, such as molecular distances between different genomic regions or between

genomic regions and distinct nuclear compartments. Until recently, studies of 3D genome folding were limited to two main technologies: imaging, particularly fluorescence in situ hybridization of DNA (DNA-FISH); and approaches based on chromosome conformation capture (3C), namely Hi-C (high-throughput chromosome conformation capture). DNA-FISH was a revolutionary approach, which allowed visualization of the spatial organization of chromosomes and genes in the nucleus^{5,6}. The approach provides single-cell information, but typically has a limited throughput that allows only a small number of genomic loci to be analysed at a time. 3C-based approaches, which depend on proximity ligation of DNA ends involved in a chromatin contact, have helped identify enhancer–promoter contacts. High-throughput derivatives, such as Hi-C, map chromatin contacts genome-wide at a length scale of hundreds of kilobases to a few megabases.

More recently, improvements in imaging techniques have increased the number of loci that can be analysed in parallel⁷ and have extended the approach to live cells^{8,9}. Orthogonal ligation-free approaches have also emerged, namely genome architecture mapping (GAM)¹⁰, split-pool recognition of interactions by tag extension (SPRITE)¹¹ and chromatin-interaction analysis via droplet-based and barcode-linked sequencing (ChIA-Drop)¹², which have started to reveal novel aspects of chromatin organization. GAM, SPRITE and ChIA-Drop map chromatin contacts genome-wide and identify topological domains but also robustly detect a previously unappreciated level of high-complexity chromatin contacts that involve three or more DNA fragments and uncover specific contacts that span tens of megabases.

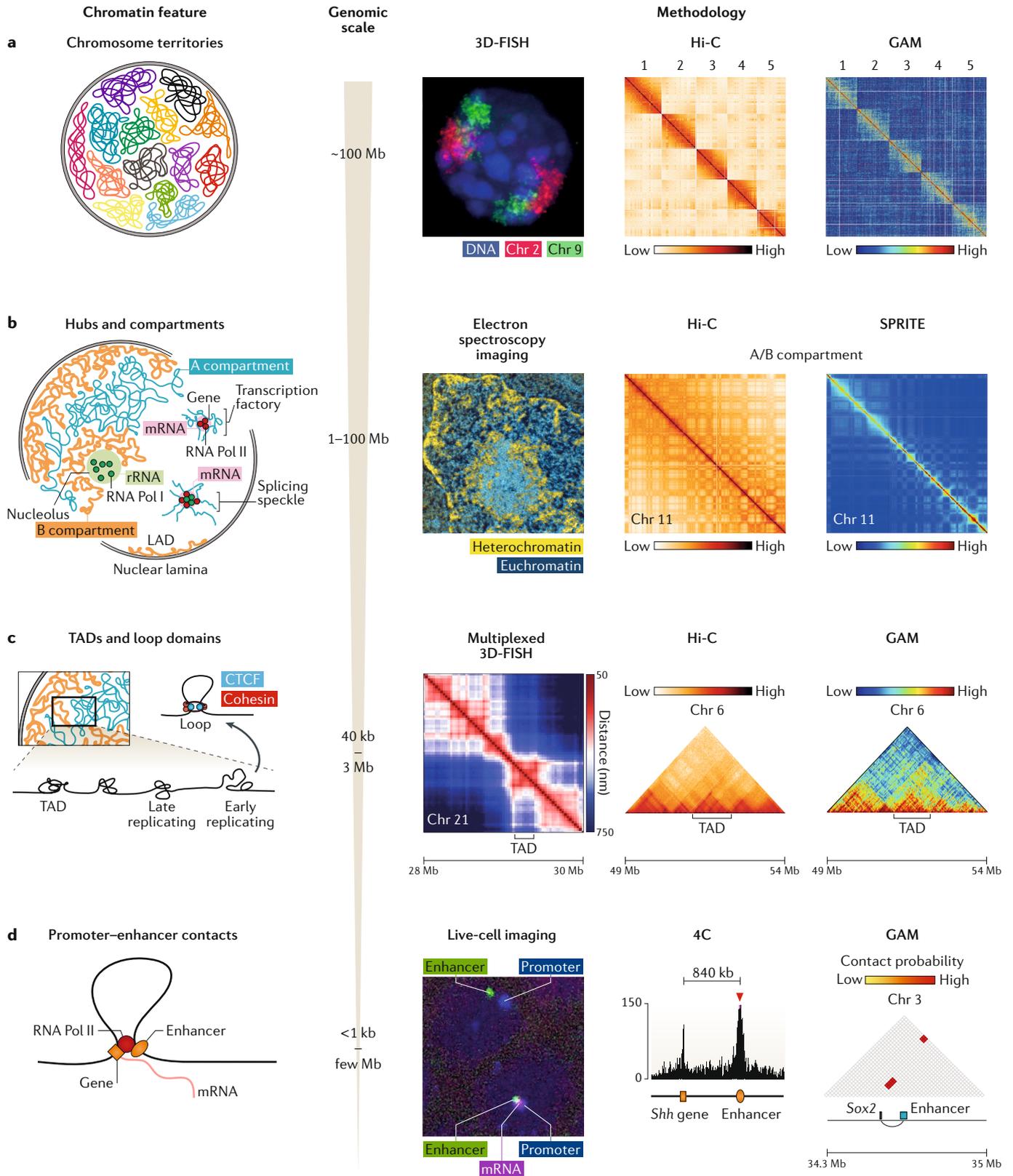
¹*Epigenetic Regulation and Chromatin Architecture Group, Berlin Institute for Medical Systems Biology, Max-Delbrück Centre for Molecular Medicine, Berlin, Germany.*

²*Institute for Biology, Humboldt University of Berlin, Berlin, Germany.*

*e-mail: rieke.kempfer@mdc-berlin.de; ana.pombo@mdc-berlin.de

<https://doi.org/10.1038/s41576-019-0195-2>

REVIEWS



Here, we review the main approaches currently used in 3D genome research, highlighting their major advantages and caveats. To recognize the strengths of each technique, it is important to understand the principles and experimental details underlying each

method, their intrinsic biases and their power to capture specific aspects of 3D genome architecture (TABLE 1). We discuss major features of 3D genome organization that have emerged, at the kilobase scale and above, through the application of these different

◀ **Fig. 1 | Methods for studying the major features of 3D chromatin folding across different genomic scales.** **a** | Chromosomes occupy discrete territories in the nucleus, which were first detected using imaging techniques. The 3D-fluorescence in situ hybridization (3D-FISH) image shows the positions of the chromosome territories of chromosome 2 (red) and chromosome 9 (green) within DAPI-stained nuclei (blue) from mouse embryonic stem cells (ESCs). Chromosome territories are also detected as regions of high-frequency intrachromosomal interactions on contact maps generated by chromosome conformation capture (3C)-based methods (such as Hi-C (high-throughput chromosome conformation capture)) and ligation-free approaches (such as genome architecture mapping (GAM)). **b** | DNA inside the nucleus separates into hubs of active (A compartment) and inactive (B compartment) chromatin, clustering around the nucleolus, splicing speckles, transcription factories and other nuclear bodies not represented here. Electron spectroscopy imaging of the mouse epiblast shows the distribution of heterochromatin (yellow) around the nucleolus (light blue) and at the nuclear periphery. Decondensed euchromatin (dark blue) is positioned more centrally in the nucleus. Nucleic acid-based structures are stained yellow, protein-based structures blue. Hi-C and split-pool recognition of interactions by tag extension (SPRITE) contact maps of mouse chromosome 11 show the separation of chromatin into discrete contact hubs (A and B compartments), which are visible as checkerboard-like contact patterns. **c** | At shorter genomic length scales, chromatin folds into topologically associating domains (TADs), which overlap with domains of early and late replication, and DNA loops, that arise from cohesin-mediated interactions between paired CCCTC-binding factor (CTCF) proteins. Multiplexed FISH of consecutive DNA segments in a 2-Mb region in the human genome shows the emergence of TADs in the population-average distance map. In Hi-C and GAM contact maps, TADs are represented by regions of high internal interaction frequencies and demarcated by a drop in local interactions at their boundaries. **d** | Contacts between a gene and its *cis*-regulatory elements occur via loop formation between the enhancer bound by RNA polymerase II (Pol II) and the gene promoter. These contacts can be detected by live-cell imaging; shown are contacts between the enhancer (green) and promoter (blue) of the *eve* gene in a *Drosophila melanogaster* embryo, with simultaneous imaging of *eve* mRNA expression (red). The circular chromosome conformation capture (4C)-sequencing track shows the interactions between the *Shh* gene promoter and the ZRS (a limb-specific enhancer of the *Shh* gene) in the anterior forelimb in mice. GAM data can be processed using the mathematic model statistical inference of co-segregation (SLICE) to extract the most significant enhancer–promoter contacts from the data set, resulting in a contact matrix with only the high-probability interactions¹⁰. The most significant interaction at the *Sox2* locus can be found between the *Sox2* gene and one of its well-studied enhancers¹⁸⁹. For parts **a**, **b** and **c**: HiGlass¹⁹⁰ was used to generate contact maps for previously published Hi-C data from mouse ESCs¹⁹¹; heat maps for GAM and SPRITE were generated from normalized published matrix files from previously published mouse ESC data (for GAM¹⁰, for SPRITE¹¹). LAD, lamina-associated domain; rRNA, ribosomal RNA. 3D-FISH image reprinted from REF.¹⁸⁷, CC BY 2.0 (<https://creativecommons.org/licenses/by/2.0/>). Electron spectroscopy image reprinted from REF.¹⁸⁸, CC BY 3.0 (<https://creativecommons.org/licenses/by/3.0/>). Part **c** reprinted with permission from REF.³², Science. Part **d** adapted from REF.⁸⁵, Springer Nature Limited, and from REF.⁴⁸, CC BY 3.0 (<https://creativecommons.org/licenses/by/3.0/>).

Topologically associating domains

(TADs). Chromosomal regions that fold into self-associating domains, with high internal interaction frequencies, demarcated by a clear drop of local interactions with neighbouring regions at their boundaries.

Chromatin loops

Local regions of high interaction frequency between two genomic loci indicate that these regions form the basis of a DNA loop. Loops often form between regions with divergent CCCTC-binding factor (CTCF) sites, or between enhancers and their target promoters.

technologies, and highlight discrepancies between approaches. We will not cover chromatin folding at the level of nucleosomes, which has been reviewed previously¹³.

Imaging-based detection of contacts

The visualization of nuclear structures and specific genomic sequences is key to understanding how chromatin is organized in the nucleus. Various light microscopy and electron microscopy techniques can be used to identify nuclear compartments or image the physical positions of specific genomic loci in the nucleus of fixed or live cells. The most commonly used imaging technique for detecting chromatin contacts in fixed cells is DNA-FISH. Contacts can be visualized in live cells using insertions of DNA binding site arrays (such as the Lac operator-repressor^{14,15}, Tet operator-repressor¹⁶

and ANCHOR¹⁷ systems) or, more recently, using CRISPR-based approaches (FIG. 2).

Measuring contacts with DNA-FISH. FISH uses fluorescently tagged DNA sequences (such as oligonucleotides) as probes to hybridize to complementary target regions of interest in the genome (FIG. 2). For hybridization to occur, a single-stranded probe needs to be able to enter the nucleus, which is usually achieved by permeabilizing the cell with a detergent or organic solvent, such as methanol. To ensure the probe can bind to its target, the DNA is most often denatured by heat and formamide treatment. The genomic regions highlighted by the hybridized fluorescent probes are then visualized by microscopy.

DNA-FISH is typically used to measure the physical distances between two or a few differentially labelled genomic regions of interest. A chromatin contact is often defined by a distance threshold, which is usually set arbitrarily according to the scale of genomic distances between the regions of interest and the resolution of the microscope. Thus, chromatin contacts have been inferred when fluorescent signals co-localize within a spatial distance of 50 nm to 1 μ m (REFS^{18–21}), although it is not clear whether distances at the top end of this range (which are close to the diameter of a whole chromosome) represent true interactions or indirect non-random positioning. DNA-FISH can also be used to visualize chromatin compaction²² or positioning of genomic regions with respect to nuclear structures, such as the nuclear lamina²³. The overall distributions of spatial distances between loci or relative to the nuclear periphery found across the cell population are usually summarized by the frequency of co-localization (that is, the frequency with which chromatin contacts are detected in the cell population), but other metrics, such as mean or median distances, are also used. The data are compared with the physical distances between other (control) genomic regions (which are often separated by similar genomic distances to the experimental loci) or, in some cases, with the nuclear diameter or volume. These metrics can help distinguish specific chromosomal conformations but can also be ambiguous, depending on the choice of control probes or if allelic differences or other forms of heterogeneity are present within the cell population.

The accuracy and power to detect different nuclear structures or contacts also depend on how well the organization of the target DNA and nuclear compartments is preserved during the FISH procedure, on the resolution of the microscope and on the size of the target genomic sequence. FISH experiments use probes made of a collection of small DNA fragments that are either synthesized (oligos) or produced by nick-translation from larger DNA molecules (plasmids, fosmids, bacterial artificial chromosomes or whole mammalian chromosomes), resulting in overlapping fragments of 100–500 bp. The probes often cover genomic sequences ranging in length from 30 kb up to entire chromosomes. The signal-to-noise ratio for locus detection increases with the target length due to increased local fluorescence and higher target specificity. Thus, with standard 3D

REVIEWS

Table 1 | Comparison of methods used to detect chromatin contacts

Assay	Description	Number of contacts per experiment	Multiplicity of contacts	Single-cell information	Number of cells	Detectable contacts	Protocol
3C-based methods							
3C	Proximity ligation and selection of target regions with primers, detection by quantitative PCR	One versus one	Pairwise	No	100 million ¹⁹²	Protein-mediated	192
4C	Proximity ligation and enrichment for contacts with one bait region by inverse PCR, detection by sequencing	One versus all	Pairwise	No	Robust: 10 million ¹⁹³ , low input: 340,000 (REF. ¹⁹⁴)	Protein-mediated	193
5C	Proximity ligation and enrichment for larger target region with primers, detection by sequencing	Many versus many	Pairwise	No	Robust: 50–70 million ¹⁹⁵ , low input: 2 million ¹⁹⁶	Protein-mediated	195,196
Hi-C	Proximity ligation and enrichment for all ligated contact pairs, detection by sequencing	All versus all	Pairwise	No	Robust: 2–5 million ⁶⁴ , low input: 100,000–500,000 (REFS ^{70,197})	Protein-mediated	64,197
TCC	Tethered proximity ligation and enrichment for all ligated contact pairs, detection by sequencing	All versus all	Pairwise	No	25 million ⁵⁷	Protein-mediated	57
PLAC-seq, ChIA-PET	Proximity ligation and pull-down of specific protein-mediated contacts, detection by sequencing	Many versus many	Pairwise	No	Robust: 100 million ¹⁹⁸ , low input: 500,000 (REF. ⁸¹)	Protein-mediated (specific)	81,198
Capture-C, C-HiC	Proximity ligation and target enrichment using probes for genomic regions of interest, detection by sequencing	Many versus all	Pairwise	No	Robust: 100,000 (REF. ¹⁹⁹), low input: 10,000–20,000 (REF. ⁹⁷)	Protein-mediated	199
Single-cell Hi-C	Proximity ligation and enrichment for all ligated contact pairs, detection by sequencing	All versus all	Pairwise	Yes	Hundreds	Protein-mediated	71
Imaging							
2D-FISH	Fixation to flatten cells, hybridization of fluorescent probes to target regions, measurement of 2D spatial distances	Between 2 and 52 regions ^a	Pairwise or more	Yes	Hundreds	All in spatial proximity	200
3D-FISH	Fixation of cells, hybridization of fluorescent probes for target regions, measurement of 3D spatial distances	Between 2 and 52 regions ^a	Pairwise or more	Yes	Hundreds	All in spatial proximity	201
Cryo-FISH	Fixation of cells, cryosectioning, hybridization of fluorescent probes for target regions, measurement of 2D spatial distances	Between 2 and 52 regions ^a	Pairwise or more	Yes	Hundreds	All in spatial proximity	101
Live-cell imaging	Fluorescent labelling of genomic loci in living cells, measurement of spatial distances over time	Between 2 and 12 regions	Pairwise or more	Yes	Hundreds	All in spatial proximity	9,39,202,203
Ligation-free methods							
GAM	Cryosectioning of fixed cells, DNA extraction from nuclear sections and sequencing, inferring spatial distances from co-segregation of genomic regions in nuclear sections	All versus all	Pairwise or more	Yes	Hundreds ¹⁰	All in spatial proximity	10
SPRITE	Fixation of cells, identification of crosslinked chromatin fragments by split-pool barcoding and sequencing	All versus all	Many	No	10 million ¹¹	Protein-mediated	11
ChIA-Drop	Fixation of cells, identification of crosslinked chromatin fragments by droplet-based and barcode-linked sequencing	All versus all	Many	No	10 million ¹²	Protein-mediated	12

3C, chromosome conformation capture; 4C, circular chromosome conformation capture; 5C, chromosome conformation capture carbon copy; ChIA-Drop, chromatin-interaction analysis via droplet-based and barcode-linked sequencing; ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; C-HiC, capture HiC; FISH, fluorescence in situ hybridization; GAM, genome architecture mapping; Hi-C, high-throughput chromosome conformation capture; PLAC-seq, proximity ligation-assisted chromatin immunoprecipitation sequencing; SPRITE, split-pool recognition of interactions by tag extension; TCC, tethered chromosome capture. ^aClassical FISH experiments rarely distinguish between more than 2–5 differentially labelled regions simultaneously²⁰⁴. Cycles of probe hybridization can increase this number up to 52 (REF.²⁰⁵).

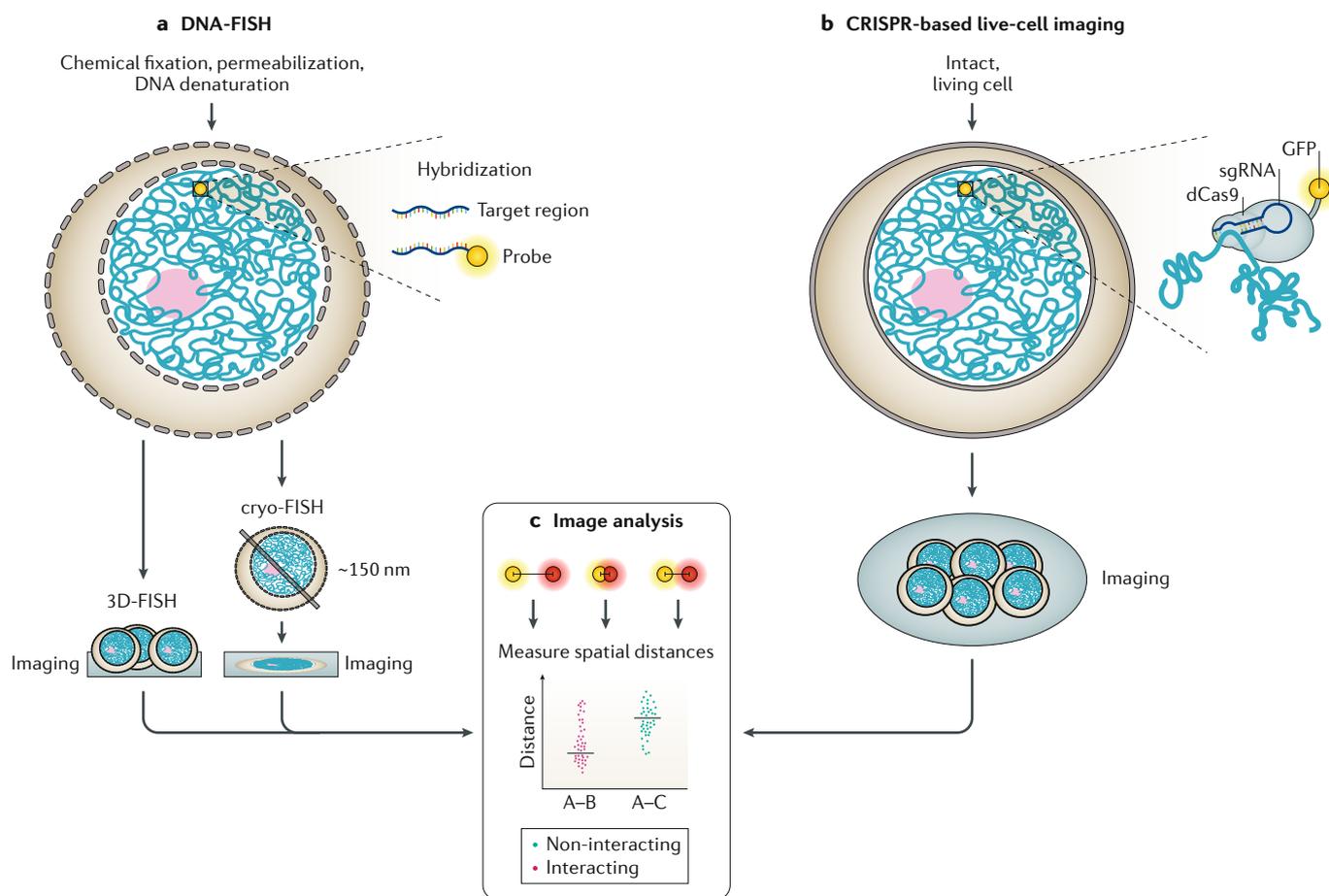


Fig. 2 | Imaging-based approaches to visualize chromatin contacts. **a** | Fluorescence in situ hybridization of DNA (DNA-FISH) uses fluorescently labelled probes that hybridize to specific genomic loci in the nucleus. Typically, cells are fixed and permeabilized and, upon denaturing of the DNA, FISH probes hybridize to their complementary target region (yellow circle). The FISH procedure can be performed in whole cells, embryos, thick tissue slices (3D-FISH) or thin cryosections of cells (cryo-FISH). The nucleolus is represented in pink. **b** | CRISPR-based live-cell imaging can be performed in the intact, living cell. Typically, dead Cas9 (dCas9) is fused to green

fluorescent protein (GFP) and the fusion protein is recruited to the target region by small guide RNAs (sgRNAs), which are complementary to the region of interest. **c** | For all of the techniques, chromatin contacts are assessed by the spatial distances between the fluorophores targeting the regions of interest (yellow and red circles). To determine the specificity of a contact, spatial distances between interacting loci should be compared with distances between non-interacting loci in numerous cells. The distribution of distances, the mean distance and the median distance can all inform about the quality and abundance of the contact in a cell population.

CCCTC-binding factor (CTCF). A transcription factor with 11 conserved zinc-finger (ZF) domains. This nuclear protein is able to use different combinations of the ZF domains to bind different DNA target sequences and proteins. CTCF is enriched at topologically associating domain (TAD) borders, where its binding can be important to specify TAD border definition.

Chromatin
The combination of DNA, RNA and protein that constitutes the chromosomes in eukaryotic cells. Broadly, heterochromatin is associated with transcriptional repression and euchromatin is associated with transcriptional activity.

FISH long-range contacts within large genomic regions, such as between TADs^{10,24} or in whole chromosomes²⁵, can be accurately detected. However, short-range interactions between chromosomal regions that are less than 100 kb apart are difficult to detect, making it harder to quantify fine-scale chromatin folding below the TAD level, such as enhancer–promoter interactions.

High-resolution imaging of chromatin contacts can be achieved using cryo-FISH, in which standard FISH probes are hybridized to thin (~100–200 nm) cryosections from cells fixed using conditions optimized to preserve the nuclear ultrastructure; the signal is then visualized using fluorescence or electron microscopy^{10,19,25–27}. More recently, the short length and high specificity of fluorophore-tagged oligonucleotides known as Oligopaints²⁸ have made it possible to target 15-kb loci using conventional microscopy²⁹ or 5-kb regions using super-resolution microscopy (when combined with a second labelling step to enhance the fluorescence signal)³⁰. Oligopaints are not derived from

cloned genomic regions but are instead generated from synthetic libraries of short (~60–100 bp) oligonucleotides, which are produced by massively parallel synthesis³¹. Once generated, the library pool can be amplified in a flexible manner, using different primer pairs to give rise to different sets of FISH probes. The ease of design of Oligopaints has opened new possibilities for the study of chromatin folding, such as being able to visualize chromatin in different epigenetic states at a resolution of tens of nanometres³². Oligopaint-based FISH has also been used in combination with high-throughput imaging to generate low-resolution contact maps (for example, at the TAD level) of whole chromosomes⁷ and high-resolution (30-kb) contact maps for stretches of DNA 1.2–2.5 Mb in length³². In addition, molecular beacon FISH probes have emerged as a way to target genomic regions as short as 2.5 kb (REF.³³). In an unbound state, these probes form a hairpin loop that minimizes the off-target fluorescent signal by bringing together the fluorescent label and a quencher.

Nuclear lamina

A protein mesh, consisting of lamins and other membrane-associated proteins, at the inner nuclear membrane that contributes to nuclear structure and function. Chromatin in the proximity of the lamina tends to be heterochromatic and transcriptionally repressed.

Fluorescence in situ hybridization

A technique that can be used to visualize the location of nucleic acid sequences within the nucleus using sequence-specific fluorescent probes that hybridize to the regions of interest, combined with microscopy.

Chromosome conformation capture

(3C). A technique used to detect the frequency of interactions between any specified two loci in the genome. Interactions between loci are captured by formaldehyde fixation, followed by restriction enzyme digestion and ligation. The frequencies of interactions between loci are determined by quantitative real-time PCR.

Hi-C

(High-throughput chromosome conformation capture). A genome-wide version of chromosome conformation capture that allows all chromatin interactions in the genome to be mapped simultaneously. The frequencies of interactions between loci are determined by paired end sequencing.

Proximity ligation

Fixation of cells, followed by fragmentation of chromatin and ligation of nearby, crosslinked DNA fragments.

Genome architecture mapping

(GAM). A genome-wide approach to detect chromatin contacts based on their physical distances within the nucleus. DNA loci are detected in thin nuclear slices by DNA extraction and sequencing. Chromatin contacts are inferred from co-segregation frequencies of pairs of DNA loci across a large (400–1,000) collection of nuclear slices.

By reducing the background signal from the unbound probe, the technique improves the visualization of small genomic regions.

Live-cell imaging of nuclear structures. Chromosome folding is a highly dynamic process that varies greatly throughout the cell cycle^{34,35}. Our ability to study these chromatin dynamics has been revolutionized by technologies based on genome editing that allow specific genomic loci to be targeted in live cells. Early iterations of this approach were rather laborious; cell lines needed to be created in which the target locus was tagged with DNA binding site arrays that recruit a fluorescently tagged cognate DNA binding protein (such as the Lac operator-repressor^{32,33}, Tet operator-repressor³⁴ and ANCHOR³⁵ systems). Now, loci can be targeted in live cells with a version of the CRISPR system that uses an endonuclease-deficient form of Cas9 (dead-Cas9 (dCas9)) fused with a fluorescent protein³⁶. The tagged dCas9 is recruited to the genomic locus of interest via its interactions with sequence-specific small guide RNAs (FIG. 2). For simultaneous labelling of two genomic regions, small guide RNAs can be differentially modified to act as scaffolds that bring fluorescent proteins to the target loci. For example, fusion proteins that comprise a fluorescent protein and either tandem dimer MS2 coat-binding protein (tdMCP) or tandem dimer PP7 coat-binding protein (tdPCP) can be directed to target loci by guide RNAs containing MS2 or PP7 aptamers, respectively. As both proteins have a comparably high exchange rate, which compensates for photobleaching, this approach is also well suited to long-term live-cell imaging^{37–39}. However, most CRISPR-based methods are currently limited to the detection of repetitive sequences because they rely on a single species of guide RNA, which hybridizes to identical genomic sequences, to direct simultaneous binding of dozens of copies of the fluorescent protein to achieve a strong fluorescent signal. A notable exception is the chimeric array of gRNA oligonucleotides (CARGO); by delivering 12 different guide RNAs into a single cell, this technique was able to efficiently label a non-repetitive 2-kb genomic region⁴⁰.

Ligation-based detection of contacts

3C-based methods extract chromatin interaction frequencies between genomic loci via chromatin cross-linking and proximity ligation (FIG. 3). Following formaldehyde fixation to capture protein-mediated and RNA-mediated contacts, chromatin is fragmented using a restriction enzyme, and the crosslinked restriction fragments are ligated⁴¹. The purified ligation fragments are called a 3C library. The ligation frequency between two loci of interest can be quantified by PCR using appropriate primer pairs. Thus, 3C focuses on interactions between two loci ('one versus one') and requires prior knowledge of the targets of interest. However, the 3C library contains all ligation products for the genome investigated and the 3C workflow can therefore be adapted to enable genome-wide analysis of chromatin contacts. Chromosome conformation capture-on-chip²⁷ or circular chromosome conformation capture⁴², both

called 4C, enrich for interactions of one region with the remaining genome ('one versus all'). Chromosome conformation capture carbon copy (5C)⁴³ captures contacts of a larger genomic stretch at high resolution ('many versus many'). Finally, Hi-C⁴⁴ captures all ligation events across the entire genome ('all versus all'). Workflows and differences between these techniques have been described elsewhere in great detail⁴⁵. Here, we focus on the most commonly used versions (FIG. 3; TABLE 1).

Mapping all contacts at a single locus with 4C. A straightforward and cost-effective method to obtain additional information from a 3C library is 4C. Here, primers for a region of interest (such as a promoter) are used to amplify all ligation partners of the locus under investigation (called the 'viewpoint') (FIG. 3). The amplified ligation products are sequenced (to a depth of 1–5 million reads per library⁴⁶) and used to analyse genome-wide interaction partners of the region of interest at a resolution of a few kilobases. 4C has been widely used to investigate *cis*-regulatory landscapes of genes, especially in development and disease⁴⁷. It is well suited for detecting short-range regulatory interactions⁴⁸, but has also been applied to detect contacts spanning long genomic distances, including whole chromosomes^{27,49}.

Mapping all contacts occurring within a large genomic region with 5C.

In 5C, large genomic regions spanning up to several megabases are amplified from the 3C library using an elegant, yet complex, mix of forward and reverse primers. For example, 5C analysis of a 4.5-Mb chromosomal region around the *Xist* gene revealed the presence of TADs²⁴. 5C has the advantage of producing high-resolution data at an affordable sequencing depth (~60 million reads per library to obtain resolution of 15–20 kb for a 1-Mb region)⁵⁰. However, the resolution of 5C is dependent on the ability to design forward and reverse primers for all possible restriction fragments across a given locus; in the absence of appropriate primers, some mappable fragments will be excluded from the contact map.

Mapping all contacts at one or more loci with capture-based methods.

A 3C library can be enriched for one or more genomic targets of interest using capture-based methods, such as Capture-C⁵¹, Capture Hi-C⁵² and CAPTURE⁵³. In these approaches, biotinylated oligonucleotides complementary to a genomic region of interest are used to pull-down specific ligation products from the library, which are then amplified and sequenced. These approaches can be used to detect interactions of one viewpoint but also of entire genomic regions⁴⁷ or groups of targets^{54,55}.

Mapping all genome-wide contacts with Hi-C and its derivatives.

Hi-C is the most commonly used genome-wide approach to map chromatin contacts from a 3C chromatin preparation⁴⁴. In this approach, the ends of crosslinked DNA restriction fragments are labelled with biotin and then ligated. After ligation, the exonuclease activity of T4 DNA polymerase is used to remove the biotin label from the ends of unligated fragments. Ligated

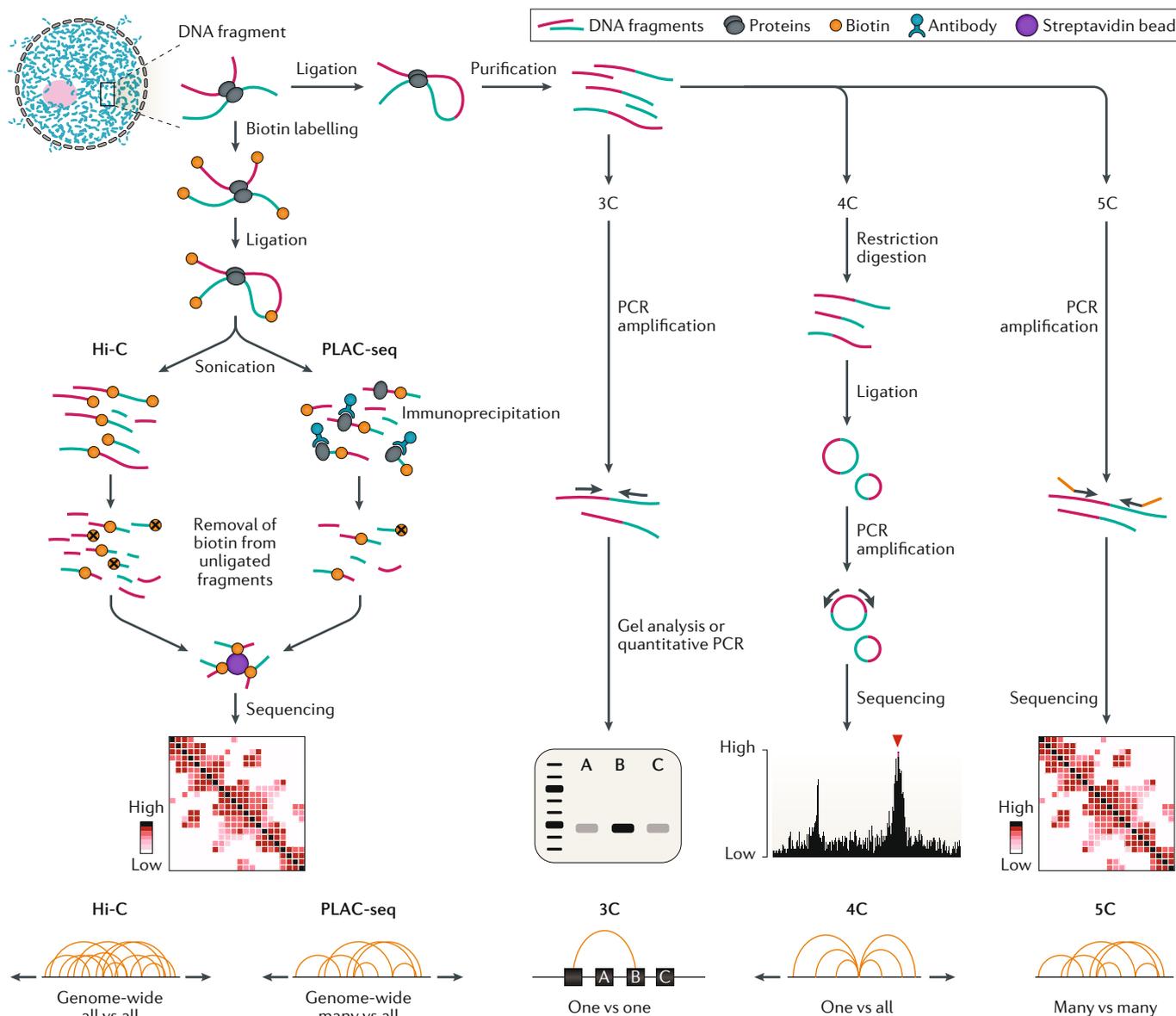


Fig. 3 | 3C and its derivatives. Chromosome conformation capture (3C)-based assays measure contact frequencies of pairs of DNA loci by proximity ligation of crosslinked and fragmented chromatin. All 3C-based assays involve fixation of the chromatin, isolation of nuclei and DNA fragmentation (for example, with a restriction enzyme). The obtained crosslinked chromatin fragments are then processed for 3C, circular chromosome conformation capture (4C) or chromosome conformation capture carbon copy (5C), which map chromatin contacts for preselected regions, or for genome-wide assays, such as high-throughput chromosome conformation capture (Hi-C) and proximity ligation-assisted chromatin immunoprecipitation sequencing (PLAC-seq). In 3C, 4C and 5C, the crosslinked chromatin fragments are ligated and the DNA is purified. In 3C, the interactions between two chosen genomic regions are detected by PCR amplification with primers specific to the two regions of interest. PCR products are analysed semi-quantitatively on an agarose gel or by real-time quantitative PCR. Interactions are defined by higher ligation frequencies compared with control regions of similar genomic distance. In 4C, interactions of one viewpoint with the whole genome are measured. The ligated and purified DNA is fractionated with a secondary restriction digest, and the digested, smaller DNA fragments are circularized and amplified with primers facing outwards from the viewpoint. The PCR products are sequenced by paired end sequencing, providing the sequence information and frequency of every chromatin contact of the viewpoint. In 5C, the ligated and purified DNA is directly amplified using primers for all restriction fragments within a consecutive genomic region, usually hundreds of kilobases up to several megabases. The PCR products are sequenced and provide information about the ligation frequencies of all fragments within the region of interest. In Hi-C and PLAC-seq, digested DNA fragments are labelled with biotin, ligated and then fragmented further by sonication. In PLAC-seq, DNA fragments bound to a protein of interest are pulled-down by immunoprecipitation. Then, in PLAC-seq and Hi-C, the DNA is purified, biotinylated nucleotides are removed from unligated fragment ends and all ligated DNA fragments are pulled-down with streptavidin beads. After pull-down, DNA fragments are sequenced and provide information about the interaction frequencies of all pairs of loci in the genome (Hi-C) or the interactions mediated by a protein of interest (PLAC-seq).

Split-pool recognition of interactions by tag extension

(SPRITE). A ligation-free approach to detect chromatin interactions by tagging crosslinked chromatin complexes. The DNA (and RNA) molecules within an individual chromatin complex are identified after sequencing by their unique combination of barcodes that have been sequentially added using a split-pool strategy.

Chromatin immunoprecipitation

(ChIP). A method used to determine whether a given protein binds to, or is localized to, specific chromatin loci in vivo, detected after (native or crosslinked) chromatin purification and immunoprecipitation, followed by DNA detection by PCR, microarray hybridization or sequencing.

fragments, which retain the biotin label, are enriched using streptavidin beads to minimize the number of unligated DNA molecules in the sequencing library. Depending on the enrichment efficiency, about 50–70% of sequencing reads map to pairs of ligated restriction fragments in Hi-C libraries⁵⁶. In tethered chromosome capture (TCC)⁵⁷, an early modification of Hi-C, the detection of unspecific ligation events between non-crosslinked material is minimized by tethering the crosslinked and biotinylated chromatin to streptavidin beads before ligation. This approach detects more long-range intrachromosomal contacts and contacts between chromosomes than standard 3C-technologies⁵⁷. By contrast, genome conformation capture (GCC)⁵⁸, an approach developed at the same time as Hi-C, sequences all DNA present in the 3C library, without preselection of ligated fragments. Although currently much more expensive, especially for large genomes, GCC has the advantage of allowing direct normalization of DNA abundance, thereby controlling for biases in sequencing and for the presence of genomic alterations, such as copy number variations. Methods for detection and normalization of copy number variations have also recently been developed for Hi-C^{59–61}.

Many other variants of genome-wide 3C-methods have been reported, ranging from technical optimizations of the original Hi-C protocol (such as DNase Hi-C^{62,63} and in situ Hi-C⁶⁴) and advances to improve resolution (such as Micro-C)^{65–67}, to protocols based on the enrichment of contacts mediated by specific proteins or open chromatin regions (open chromatin enrichment and network Hi-C (OCEAN-C)⁶⁸). Currently, the most commonly used version is in situ Hi-C. In the original Hi-C protocol, sodium dodecyl sulfate (SDS) is used to disrupt the nuclear membrane and ligation of crosslinked DNA therefore occurs partially in solution. In situ Hi-C omits this SDS step, allowing ligation of chromatin fragments within the presumably more native environment of the intact nucleus. As a result, the number of random ligation events is reduced and signal-to-noise ratios are improved, thereby reducing the sequencing depth and enabling higher-resolution contact maps. However, detailed analyses of the nuclear fragments that contribute to contacts in the original version of Hi-C showed that large portions of the chromatin were thought to remain inside the partially digested nucleus during ligation⁶⁹. Nonetheless, the in situ Hi-C protocol is faster and easier than the original version⁶⁴, mainly because it does not require extensive dilution of the crosslinked chromatin prior to DNA ligation. Consequently, all subsequent steps can be conducted in smaller volumes, allowing more efficient ligation and DNA extraction. Easy Hi-C is another recent approach to simplify Hi-C⁷⁰. It avoids biotin enrichment and can be used with lower cell numbers than standard Hi-C (TABLE 1).

Mapping genome-wide contacts in single cells with single-cell Hi-C. Standard Hi-C generates average contact maps from millions of cells, without any possibility to understand heterogeneity of the cell population. Single-cell Hi-C overcomes this limitation by allowing Hi-C contact maps to be produced from individual cells isolated during the process of generating Hi-C libraries^{71,72}.

This approach allows rare cell types to be studied⁷³ and helps chromosome structures to be determined at specific stages of the cell cycle⁷⁴. The single-cell Hi-C protocol involves in situ proximity ligation of crosslinked and digested chromatin, followed by isolation of single nuclei from the cell suspension and generation of sequencing libraries from each nucleus^{71,74}. Single-cell combinatorial indexed Hi-C (sciHi-C) adopts a different approach; instead of isolating single cells, DNA within each nucleus is tagged with a unique combination of barcodes⁷⁵. First, cells are fixed, lysed and digested with a restriction enzyme. Then, the cell suspension of digested, but intact, nuclei is split into 96-well plates, indexed with individual barcodes, pooled and split again. After several rounds of indexing, in situ proximity ligation and library preparation are performed on pooled nuclei, allowing high-throughput generation of single-cell Hi-C libraries.

One of the major challenges in single-cell Hi-C is the efficient recovery of contacts: inefficient digestion and ligation and incomplete recovery of input material result in contact maps that represent only a proportion of the contacts that may exist in a single cell. Modifications of the original protocol increased the average number of contacts detected in one cell from ten thousands up to hundreds of thousands^{34,74}, but this remained a fraction (2–5%) of the possible contacts in the genome. Recently, the development of Dip-C (diploid chromatin conformation capture) has increased the number of detectable contacts to an average of 1 million per cell by omitting biotin incorporation and including a whole-genome amplification step in the protocol⁷⁶.

Combining 3C-based approaches with chromatin immunoprecipitation.

3C-based methods can be used to study chromatin contacts mediated by specific proteins, such as chromatin modifiers, architectural proteins, members of the transcription machinery or cell type-specific transcription factors. To explore contacts that coincide with chromatin occupancy of specific proteins, Hi-C libraries can be enriched by chromatin immunoprecipitation (ChIP) before ligation. Early methods, such as ChIP-loop⁷⁷ and enhanced 4C-ChIP (e4C)⁷⁸, required that chromatin be solubilized to enable specific immunoprecipitation before ligation. However, standard 3C conditions often do not fully solubilize chromatin, as nuclei stay mostly intact after SDS treatment⁶⁹, resulting in low signal-to-noise ratios. Other approaches, such as chromatin interaction analysis by paired-end tag sequencing (ChIA-PET), included sonication of the nuclei, as is more typically used for ChIP⁷⁹. Although sonication allows efficient precipitation of chromatin, its influence on the outcome of the subsequent proximity ligation remains unclear. Challenges in implementing ChIA-PET have led to other strategies for combining ChIP with Hi-C, namely Hi-ChIP⁸⁰ and proximity ligation-assisted chromatin immunoprecipitation sequencing (PLAC-seq)⁸¹. Instead of performing protein pull-down followed by ligation of DNA fragments, Hi-ChIP and PLAC-seq perform in situ Hi-C and proximity ligation before sonication and immunoprecipitation. In this order, the ligation occurs in intact nuclei under optimal conditions, before chromatin contacts specific to the protein of interest are

enriched. Regardless of these increased efficiencies, the results from immunoprecipitated 3C-libraries should be interpreted carefully because of the bias introduced by enriching for genomic regions that are bound by the protein of interest⁸².

Genomic resolution of genome-wide 3C-methods. A major consideration for any genome-wide technique is genomic resolution. Hi-C data represent interaction frequencies between genomic regions in a contact matrix, consisting of equally sized genomic bins. The bin size (resolution) depends almost entirely on the sequencing depth. Resolutions of 30 kb or lower are often preferred to study the chromatin domain and compartments but also long-range contacts between large genomic regions (such as TADs); using standard Hi-C, this requires sequencing depths of approximately 200–400 million reads in mammalian genomes. However, billions of reads become necessary for high-resolution (1-kb) data sets of the human genome that can provide detailed insights into 3D genome topology⁶⁴. Recently, a computational approach, HiCPlus, applied deep learning to infer high-resolution contact matrices from low-resolution Hi-C data, which reduced the sequencing depth required to obtain a given resolution by a factor of 16 (REF.⁸³).

Ligation-free detection of contacts

The reliance of 3C-based approaches on the ligation of the ends of DNA fragments found in a cluster of contacts favours the detection of ‘simple’ chromatin contacts which involve two or a few genomic regions. This bias occurs because each DNA fragment can ligate with only one or two other fragments, so not all instances of every interaction in a complex cluster are detected⁸⁴. Thus, the full interactome of each DNA fragment is diluted by the choice of only one or two other fragments during ligation. Recently, three ligation-free approaches have been developed for genome-wide mapping of chromatin contacts: GAM¹⁰, SPRITE¹¹ and ChIA-Drop¹². These methods are orthogonal to ligation-based approaches and are starting to provide new insights into 3D genome topology. Other ligation-free approaches — tyramide signal amplification (TSA-seq)⁸⁵ and DNA adenine methyltransferase identification (DamID)^{86–88} — map chromatin with respect to nuclear landmarks (such as the nuclear lamina or various nuclear bodies), thereby helping to define chromatin positions in 3D space.

Mapping contacts with nuclear structures with DamID and TSA-seq. DamID is an *in vivo* genome-wide method for detecting interaction sites between a protein of interest and DNA. The DNA binding domain of the protein of interest (for example, RNA polymerase II (Pol II)) is fused to the DNA adenine methyltransferase (Dam) protein from *Escherichia coli*^{86–88}, which specifically methylates adenines in the sequence GATC. When the fusion protein is expressed at low levels in cells, GATC sequences within or close to DNA binding sites of the protein of interest are marked by methylation. After DNA extraction, the methylated GATC sites are cut with a methylation-sensitive restriction enzyme and adapters are added to the restriction fragments to

ensure only methylated binding sites are amplified and sequenced. In an interesting adaptation called targeted DamID (TaDa)⁸⁸, expression of the Dam fusion protein is restricted to a specific cell type of interest, using targeted expression systems (such as the Gal4–UAS system), which allows detection of DNA–protein interactions, in a cell type-specific manner without prior isolation or sorting of cells. DamID has been successfully used to study DNA interactions with proteins such as Lamin B1, which resulted in the genome-wide mapping of lamina-associated domains and provided spatial information about chromatin with regard to the nuclear periphery^{89,90}. However, interactions between chromatin and other nuclear compartments, such as splicing speckles, are not readily detected with DamID because most of the DNA surrounding these compartments does not directly bind to the tagged proteins⁹¹.

TSA-seq addresses this problem using tyramide signal amplification to measure the distances between chromatin and nuclear compartments⁹². In this approach, horseradish peroxidase (HRP) is conjugated to an antibody that binds to a protein specific to the nuclear compartment of interest, where it catalyses the production of biotin-conjugated tyramide free radicals, which diffuse and bind to nearby macromolecules — including DNA. Biotin-labelled DNA can be subsequently selected by biotin pull-down and sequenced to identify all genomic regions that were close enough to the protein of interest to be labelled. TSA-seq has been used to map genome-wide the distances between all genes and their nearest splicing speckle⁹².

Another recent adaptation of DamID, called DamC, detects 4C-like contacts between a target region and the surrounding DNA regions, up to distances of a few hundred kilobases⁹³. In DamC, Dam is fused with the reverse tetracycline receptor (rTetR), which binds to Tet operator sites inserted at the genomic region of interest. The Dam fusion protein methylates the target and its interaction partners *in vivo*. When combined with high-throughput sequencing, DamC reveals chromatin contacts independently of crosslinking or ligation, but unlike the other 3C-methods and ligation-free approaches it requires engineering of the cells of interest. Comparison of DamC data with 4C and Hi-C data showed high similarities at the level of TADs and CTCF loops at many genomic sites; however, some differences at loops and sub-TAD structures could also be observed⁹³.

Mapping all genome-wide contacts with GAM. In GAM, nuclei are sectioned in random orientations from a population of fixed and sucrose-embedded cells using ultra-thin cryosectioning (220 nm thickness). Single nuclear slices are then isolated directly from the cryosection by laser microdissection. GAM thus avoids cell extraction or sorting, both of which can disrupt cellular and nuclear structures, which can be especially important when analysing complex tissues. The DNA from every slice is extracted, whole-genome amplification is performed and indexed sequencing adapters are added before the DNA from all slices is pooled for sequencing (FIG. 4). From the sequencing data for several hundred nuclear sections, each from a single cell,

Genomic resolution

The size of the window (often in the range of kilobases) when, for most assays, reads after sequencing are mapped to the genome and then binned into equally sized genomic windows (bins).

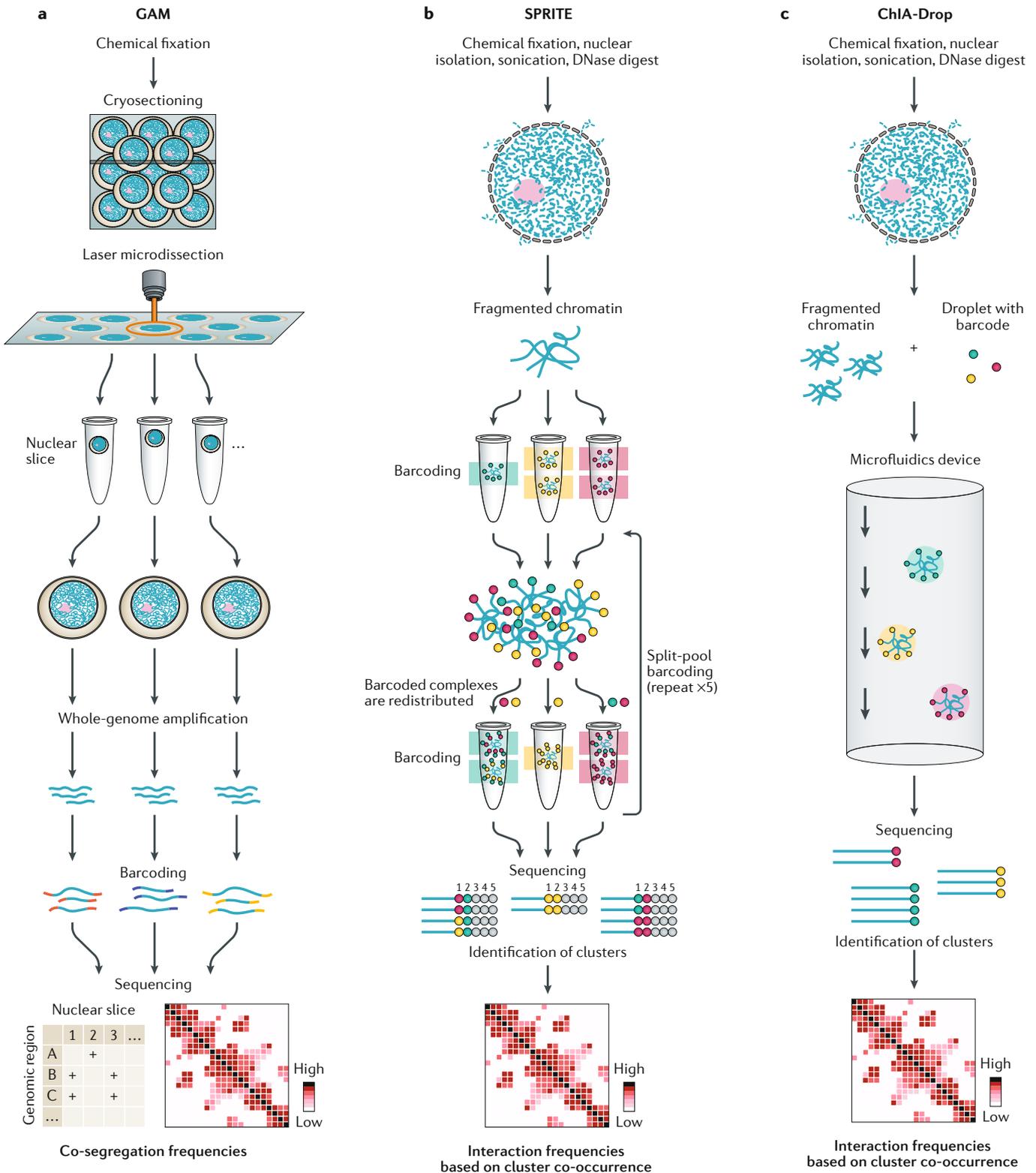
Sequencing depth

The average number of reads representing a given nucleotide in the reconstructed sequence. A 10× sequence depth means that each nucleotide of the transcript was sequenced, on average, 10 times.

Nuclear bodies

Membrane-less compartments in the nucleus with high concentrations of DNA binding proteins, chromatin modifiers or RNAs that can be involved in shaping chromatin structure and modulating gene regulation. Nuclear bodies include the nucleolus, splicing speckles and Polycomb bodies.

REVIEWS



chromatin contacts between pairs of DNA loci can be inferred by counting their co-segregation frequency (that is, how often the two loci are contained in the same nuclear sections). Genomic regions that are closer in 3D space are more frequently found in the same nuclear slice. To detect statistically significant interactions, GAM was combined with a mathematical

model, statistical inference of co-segregation (SLICE)¹⁰. The most specific chromatin contacts detected with SLICE were found to contain active genomic regions, such as active enhancers and actively transcribed genes, with these contacts extending over megabases up to entire chromosomes¹⁰. SLICE separately models the random interactions that depend on genomic distance and

◀ **Fig. 4 | Ligation-free methods to map chromatin contacts genome-wide.** **a** | Genome architecture mapping (GAM) measures co-segregation frequencies of genomic regions by slicing the nucleus into thin nuclear sections and sequencing the DNA content of a large number of randomly collected slices. To obtain nuclear slices, cells are fixed and cryosectioned. Single nuclear slices are isolated from the cryosection using laser microdissection. DNA is extracted from each nuclear slice by whole-genome amplification and sequenced. The sequence information is used to score the presence or absence of genomic loci in each slice. Spatial proximity of all pairs of loci in the genome is inferred from the frequency of their co-occurrence in the population of slices. **b** | Split-pool recognition of interactions by tag extension (SPRITE) detects chromatin interactions of multiple genomic regions by tagging single crosslinked chromatin complexes with unique combinations of identifiers before sequencing. Cells are fixed and the crosslinked chromatin is fragmented using sonication. The resulting chromatin complexes are split into wells of a 96-well plate, and DNA in each well is ligated to a unique barcoded adapter. The contents from all wells are pooled and split again, followed by adapter ligation. The process is repeated five times so that each chromatin complex is labelled with a unique combination of adapter sequences. DNA is purified and sequenced, and the adapter combination of each sequenced DNA fragment is used to identify all genomic regions that share the same combination of adapters and were, therefore, initially crosslinked together, inferring spatial proximity. **c** | Chromatin-interaction analysis via droplet-based and barcode-linked sequencing (ChIA-Drop) detects chromatin contacts by barcoding crosslinked chromatin complexes after cell fixation, lysis and chromatin fragmentation. Barcodes are delivered in a droplet that contains a unique identifier and reactions for adapter ligation and DNA amplification. Each chromatin complex is loaded onto a droplet in a microfluidics device and sequenced. Barcodes identify regions from the same droplet, indicating regions that were crosslinked due to spatial proximity.

the specific interactions that occur at a given physical distance (for example, below 100 nm)¹⁰; it interrogates which pairs of loci co-segregate more often in the collection of slices than expected from random contacts, and quantifies the frequency of specific interaction in the cell population. GAM also allows genome-wide interactions between three or more DNA loci to be detected simultaneously, and has detected long-range contacts between TADs containing super-enhancers and highly-transcribed TADs¹⁰. The resolution of GAM data sets depends on the number of nuclear slices collected. With 400 nuclear slices, sequenced with ~1 million reads per slice, it was possible to achieve a resolution of 30 kb for pairwise chromatin contacts¹⁰, comparable with a Hi-C library with similar sequencing depth⁹⁴. Larger GAM data sets comprising a few thousand nuclear slices will help define the maximal resolution that can be practically afforded by GAM.

Mapping all genome-wide contacts with SPRITE and ChIA-Drop. SPRITE¹¹ and ChIA-Drop¹² detect chromatin interactions by tagging crosslinked chromatin complexes. Similar to 3C-based approaches, these methods rely on mild fixation and fragmentation of chromatin inside the nucleus – but unlike 3C-based approaches, they do not use proximity ligation. Instead, in SPRITE, the crosslinked chromatin fragments are split across a 96-well plate, where each well contains a unique barcode (FIG. 4). The indexed chromatin complexes are re-pooled, followed by sequential rounds of splitting, barcoding and pooling. The DNA (and RNA) molecules within an individual chromatin complex are identified after sequencing by their unique combination of barcodes added using this split-pool strategy; only DNA fragments that were crosslinked with each other will display the same combinations of barcodes. In ChIA-Drop, crosslinked and fragmented chromatin is separated into

single chromatin complexes by droplet formation using a microfluidics device. Each droplet contains reagents for barcoding and amplification, and barcoded complexes are pooled and sequenced, as in SPRITE. SPRITE detects TADs and loop domains, both of which are features of Hi-C contact maps. However, SPRITE also detects additional genome-wide features of nuclear architecture, such as the association of specific genomic regions with nucleoli and splicing speckles. The predominant chromatin hubs around these nuclear bodies contain genomic regions from different chromosomes, an observation that is in agreement with single-cell imaging⁹⁵ but that had not been made using 3C-based assays. SPRITE also detects long-range contacts between regions containing active genes and super-enhancer regions that were first recognized as being multiway-specific interactions in a study using GAM¹⁰.

Comparing approaches

Fundamental differences exist between current approaches for mapping 3D genome folding, including how the chromatin is fixed and prepared, their power to detect multiple chromatin contacts or contacts with different spatial distances and protein occupancy, and their ability to detect long-range contacts within the same (FIG. 5) or different chromosomes. These differences have sometimes led to observations that can be difficult to reconcile between the different approaches.

Fixation and chromatin preparation. With the exception of live-cell methods (such as DAM-based and CRISPR-based approaches), all chromatin folding techniques start by crosslinking DNA–protein complexes to stabilize nuclear structures (TABLE 2). Chemical fixation using formaldehyde is the most common approach for crosslinking, but concentrations, buffers and fixation times vary widely; for example, 1% formaldehyde is typically used for 3C-based methods, 4% for most DNA-FISH experiments in whole cells and 8% for GAM or cryo-FISH in nuclear slices. Other fixatives include solvent-based precipitation using ethanol, methanol or acetone. A recent imaging study compared the effects of formaldehyde fixation and cryofixation on nuclear structure using partial wave spectroscopy⁹⁶. It revealed that weaker fixatives (such as 4% formaldehyde in PBS) introduce larger structural distortions than stronger fixatives (such as glutaraldehyde, often used for electron microscopy). However, the distinction between condensed and decondensed chromatin remains detectable at the population level⁹⁶, which is consistent with the ability of all current chromatin folding methods to successfully map euchromatin and heterochromatin. The effect of varying crosslinking conditions (from no fixation to 5% formaldehyde fixation) has been examined in Capture-C experiments; similar short-range interactions were detected under all conditions, but formaldehyde concentrations below 2% improved the efficiency of detection⁹⁷. Our own previous work showed that the organization of the active form of Pol II, which marks transcription sites, can be highly disrupted with weaker fixatives, but not with the fixation regimen used for GAM or cryo-FISH⁹⁸. In FISH, denaturation of the DNA

REVIEWS

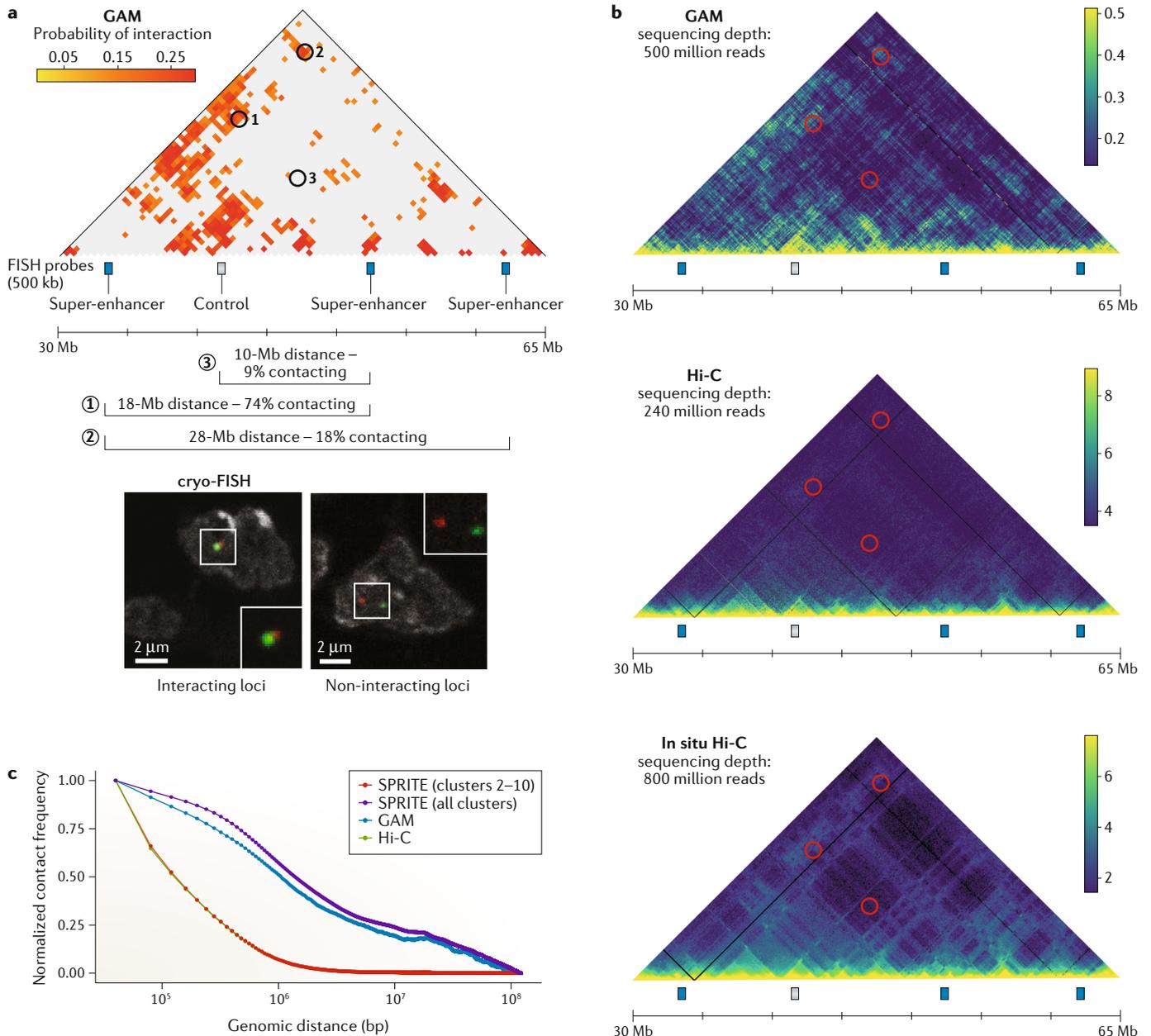


Fig. 5 | Comparison of long-range chromatin contacts across methods.

a | Genome architecture mapping (GAM) detects significant interactions between super-enhancers (circles 1 and 2) that span large genomic distances (18 Mb and 28 Mb). The heat map shows GAM interaction probabilities for chromosome 11 (region 30–65 Mb) in mouse embryonic stem cells (ESCs) at 500-kb resolution¹⁰. Contacts between the super-enhancer regions can also be detected using cryo-fluorescence in situ hybridization (cryo-FISH), which confirmed their interactions in a high percentage of cells in the population: the 18-Mb distant super-enhancer contact and the 28-Mb distant contact were detected in 74% and 18% of cells, respectively. By comparison, contacts between a super-enhancer and a 10-Mb distant control region (circle 3) not detected by GAM were detected in only 9% of cells. The images show DAPI-stained cryosections with interacting and non-interacting 500-kb FISH probes.

b | Long-range super-enhancer contacts can also be found when looking at GAM contacts without filtering for the most significant interactions (~500 million reads, 40-kb resolution, mouse ESCs, data from REF.¹⁰), and although they are not readily detected in Hi-C data with an average sequencing depth (~240 million reads, 40-kb resolution, mouse ESCs, data from REF.⁹⁴), they start to emerge in deep-sequenced in situ Hi-C data (~800 million reads, 50-kb resolution, mouse ESCs, data from REF.¹⁹¹). Heat maps were generated by Christophe Thieme from the published, normalized matrix files.

Scores are colour-coded, where the colour-code range (maximum and minimum cutoffs) is determined by the mean value of the bin distances 1–20 and –50 to –30 from the diagonal, respectively. **c** | The plot shows the distribution of contact frequencies detected by high-throughput chromosome conformation capture (Hi-C), GAM and split-pool recognition of interactions by tag extension (SPRITE) (all clusters or clusters with 2–10 reads) along the linear genomic distance of chromosome 11 in mouse ESCs, scaled to the maximum observed value in each data set. The ligation-free methods GAM and SPRITE detect similar ranges of chromatin contacts, which can extend over large genomic distances. By contrast, Hi-C contacts typically extend over shorter genomic distances. However, SPRITE data can be sorted based on the number of interactions within one chromatin complex. When considering only small SPRITE clusters with fewer than 10 genomic regions in the same chromatin cluster, the range of detection between Hi-C and SPRITE is comparable, indicating that Hi-C favours less complex short-range contacts over long-range interactions involved in chromatin hubs with many interaction partners. The plot was generated using the same data for GAM (REF.¹⁰) and Hi-C (REF.⁹⁴) as used in part **b**, and data provided by Sofia Quinodoz for normalized SPRITE clusters for chromosome 11, according to figure 3B of REF.¹¹. Part **a** adapted from REF.¹⁰, Springer Nature Limited. Heat maps in part **b** and plot in part **c** courtesy of Christoph Thieme, Max Delbrueck Center, Germany.

Table 2 | Experimental differences between chromatin contact assays and their effects on nuclear structures

Assay	Fixation	Chromatin preparation	Effects on chromatin
2D-FISH	Hypotonic treatment, followed by methanol–acetic acid for 10 min	Permeabilization, HCl treatment, heat and formamide denaturation, physical flattening of cells through gravity and drying	The fixation process flattens cells, alters the shape of the nucleus and the structure of chromocentres, and changes the association of some loci with their chromosome territories ^{206–209} ; denaturation can lead to loss of DNA ²¹⁰
3D-FISH	2–4% depolymerized PFA for 10 min (fixatives are buffered with PBS)	Permeabilization, HCl treatment, heat and formamide denaturation	Fixation can cause nuclei to increase in size, accompanied by a change in the positions of distant chromatin domains; overall, nuclear structures and the distribution of chromatin in the nucleus remain intact; denaturation disrupts the nuclear membrane, accompanied by loss of some heterochromatic regions and redistribution of histone H2B ⁶⁹
Cryo-FISH and GAM	4% depolymerized PFA for 10 min, followed by 8% PFA for 2 h (fixatives are buffered with 0.25 M HEPES)	Sucrose-embedding for cryoprotection, cryosectioning. For cryo-FISH, slices are permeabilized, treated with HCl, heat and formamide denaturation	Stringent, electron microscopy-grade fixation preserves nuclear and cytoplasmic ultrastructures and is indistinguishable from glutaraldehyde fixation ⁹⁸ ; robust fixation renders nuclear structures more resilient to the negative effects of heat denaturation for FISH, with good preservation of nuclear compartments before and after FISH ^{25,101}
3C-based methods, SPRITE and ChIA-Drop	1% formaldehyde for 10 min (fixatives are buffered with cell culture medium or PBS)	Cell lysis, SDS treatment, fragmentation of the genome (for example, restriction digest, DNase treatment, sonication)	Mild fixation is required to allow subsequent restriction digest, but can result in redistribution of nuclear proteins (similar fixation studied in REF. ⁹⁸); after chromatin preparation, restriction enzymes disrupt nuclear structures, nuclei swell and chromatin is distributed more uniformly ⁶⁹ ; DNA-FISH on digested nuclei shows that chromosome territories seem larger and less condensed compared with fixed cells before digestion, but DNA is maintained in territories ⁶⁹

3C, chromosome conformation capture; ChIA-Drop, chromatin-interaction analysis via droplet-based and barcode-linked sequencing; DNA-FISH, fluorescence in situ hybridization of DNA; FISH, fluorescence in situ hybridization; GAM, genome architecture mapping; HCl, hydrogen chloride; PFA, paraformaldehyde; SDS, sodium dodecyl sulfate; SPRITE, split-pool recognition of interactions by tag extension.

by heat and formamide induces fine structural changes in chromatin folding, such as slight distortions of the interchromatin space^{99,100}. However, 3D-FISH preserves the organization of centromeres seen by imaging the same cells before and after hybridization¹⁰⁰, and cryo-FISH retains the organization of active Pol II sites¹⁰¹. In another method, resolution after single-strand exonuclease resection (RASER)-FISH, heat denaturation of the DNA is avoided and DNA accessibility is achieved by exonuclease digestion, thereby reducing the effects of DNA denaturation¹⁰².

Multiplicity of chromatin contacts. The dependency of 3C-based methods on DNA end ligation results in preferential detection of low-multiplicity contacts that involve only a few genomic regions⁸⁴. However, every 3C library also includes interaction events that occur at complex clusters involving more than two DNA fragments, albeit at a low representation. Current methods to capture these higher-complexity ligation events include multi-contact 4C (MC-4C)¹⁰³, which uses long-read sequencing (such as nanopore sequencing) of 4C libraries to capture three-way contacts of a region of interest, and chromosomal walks (C-walks)¹⁰⁴, which implement multiple ligation steps followed by dilution and barcoding of the isolated ligation products. Alternatively, methods such as the concatemer ligation assay (COLA)¹⁰⁵ and Tri-C¹⁰⁶ generate 3C libraries with a restriction enzyme

that cuts small DNA fragments, which increases the frequency of detecting multiple ligation events in one sequencing read. Estimates based on direct comparison of pairwise and multiway ligation events indicate that only 17% of chromatin contacts in mouse embryonic stem cells (ESCs) are pairwise contacts and, therefore, the majority of the genome is involved in higher-order contacts between more than two genomic loci¹⁰⁴. These observations are supported by a recent study using SPRITE, which showed that classical ligation-dependent methods under-represent higher complexity contacts¹¹ (FIG. 5c). Assays that do not depend on ligation detect DNA fragments that are in spatial proximity regardless of the number of interacting genomic loci. For example, long-range multiway contacts between genomic regions harbouring super-enhancers were readily found in GAM¹⁰, FISH¹⁰ and SPRITE¹¹ data, but had not previously been detected with Hi-C. Furthermore, analyses of triplet interactions between TADs in GAM showed that multiple interactions between super-enhancer regions and active genes are a common feature of genome conformation in mouse ESCs¹⁰.

Spatial distance between contacting genomic regions. The spatial distance between genomic loci is thought to influence the probability of ligation irrespective of the frequency of contacts. Whereas cryo-FISH and SPRITE have readily detected abundant interchromosomal

contacts in human, mouse and *Drosophila* cells^{25,76,107,108}, 3C-based methodologies are more often used to explore specific contacts within chromosomes, with some exceptions^{27,76,109–111}. Recent CRISPR–Cas9 live-cell imaging of a small number of chromatin contacts within and between chromosomes showed that interchromosomal contacts display spatial distances in the range of ~280 nm, in contrast to distances of ~190 nm for intrachromosomal interactions²⁰. Interestingly, only the intrachromosomal contacts could be observed in matching Hi-C data, indicating a dependency of close spatial distances for successful proximity ligation.

Protein-mediated interactions versus bystander contacts. GAM and all imaging-based techniques collect all possible spatial relationships between genomic regions, regardless of their involvement in a protein-mediated interaction, and allow sampling of the whole range of spatial distances within the interphase nucleus. Thus, these methods also detect bystander contacts. However, it is possible to identify the most specific contacts through effective sampling to take into account all behaviours of all genomic regions at all linear distances across the cell population. In this regard, GAM currently has more statistical power than FISH as it samples all possible combinations, whereas FISH remains limited to the analyses of a subset of regions or chromosomes.

Levels of concordance between different methods. The validation of results obtained by 3C-based methods often entails the use of DNA-FISH on a few selected loci. Many examples show agreement between 3C interaction frequencies and spatial distances measured by FISH, especially at large genomic distances^{44,64,112–115}. Loci in the same TAD are often closer in nuclear distance than loci in different TADs^{24,94}, and interaction frequencies obtained from Hi-C correlate with spatial distances at and above the TAD level¹¹⁵. A linear relationship between Hi-C contacts and FISH distances was found by investigating the physical distances between all TADs along a chromosome¹¹⁵. An overall correlation between Hi-C interactions and the median spatial distance measured by high-throughput FISH have recently been shown for 90 pairs of loci. However, the range of physical distances between genomic regions containing Hi-C interactors (with high ligation frequency) and non-interactors (with low ligation frequency) overlap extensively, with about 20% of distances being closer between two non-interactors than two interactors²¹. Thus, Hi-C captures spatial proximity but Hi-C interactions are not easily translated into physical distances. Other comparisons between FISH and 3C-based methods have also found non-trivial relationships between physical distance distributions and population-average interaction frequencies¹¹³ and show that contact frequency is distinct from average spatial distance, both in polymer simulations and in experimental data¹¹⁶.

The use of FISH to validate Hi-C results has helped investigate false positives in Hi-C data, assuming FISH is correct, but is not a valid strategy for an unbiased search for contacts that are missed by Hi-C (that is, false negatives). Thus, any under-represented contacts in Hi-C

data have so far not been systematically studied. The development of orthogonal genome-wide ligation-free approaches, such as GAM and SPRITE, have been able to identify new aspects of 3D genome folding that had not been detected by Hi-C but which are fully validated by FISH^{10,11}. The first and relatively small GAM data set, combined with the mathematical model SLICE, identified specific long-range contacts across genomic distances that span tens of megabases, which involve active and enhancer-rich genomic regions (FIG. 5a,b). One promising outcome of the emergence of these orthogonal approaches is the development of analysis tools that use the information they generate about such long-range contacts to discover the same contacts in Hi-C data. In this regard, it is interesting to note that CTCF depletion in human cells results in the detection by Hi-C of long-range contacts between super-enhancers¹¹⁷, which raises the possibility that CTCF-mediated contacts may be preferentially detected by Hi-C in normal conditions, but once CTCF-dependent interactions are lost, other underlying folding patterns, including long-range contacts, become easier to detect.

The first SPRITE data set has also highlighted novel aspects of 3D folding that are not readily captured by Hi-C¹¹. By discriminating contacts according to their multiplicity, SPRITE shows a contact decay with genomic distance that is Chromosome territories and interchromosomal very similar to Hi-C when considering only low-complexity SPRITE clusters (2–10 genomic regions per contact hub; FIG. 5c). By contrast, SPRITE shows a striking abundance of long-range contacts when considering also higher-order contacts, which confirms early theoretical predictions that ligation-based approaches are biased to the detection of more simple 3D chromatin contacts⁸⁴. Although GAM and SPRITE are orthogonal methodologies, their frequency of contacts relative to genomic distance are remarkably concordant^{10,11} (FIG. 5c).

Limitations and applications of different methodologies. Methods that use proximity ligation are limited by the low efficiency of ligation, and are also potentially affected by the local distance between, or the topology of, the two DNA ends within the cluster of contacting DNA fragments. SPRITE also depends on ligation of a small oligo to each DNA end in a contact cluster; however, it is no longer dependent on the physical distance between two DNA fragments in the cluster, which allows mapping of all contacts within one chromatin complex. In 3C-based methods and SPRITE, detection of contacts depends on the efficiency of the fragmentation step to expose the DNA end. In GAM, there is no DNA restriction digest or ligation, and the detection of DNA depends on its extractability and sequencing depth. 3C-based methods, GAM, SPRITE and FISH can be applied directly to cells, tissues or organisms, whereas insertions of DNA binding site arrays (such as the Lac operator system), CRISPR-based imaging and Dam-related methods require genetic engineering of cell lines or whole organisms, and will not be suitable for the analyses of most human biopsies.

Each of the assays discussed here has different limitations and applications, and thus contributes to our

current understanding of 3D genome folding in different ways. 3C-based techniques have the advantage of providing enormous amounts of chromatin contact information in one comparably simple biochemical experiment, although they may require high-depth sequencing when aiming for high resolution. 3C-based methods, and in particular proximity ligation itself, also have important limitations that favour the detection of more simple contacts over higher-order chromatin contacts, which can lead to misunderstanding the importance and abundance of certain interactions. However, 3C-based techniques are well suited for studying local chromatin folding within the range of kilobases up to a few megabases.

Imaging and ligation-free methods have the ability to detect chromatin contacts at all scales of chromosome folding, including contacts between chromosomes. GAM and SPRITE can be readily used for sequence-unbiased genome-wide explorations, whereas detection of contacts with DNA-FISH remains limited to preselected loci and is most often used to validate findings from genome-wide techniques. Imaging fluorescently labelled chromatin loci in live cells with CRISPR-based techniques will improve our understanding of possible artefacts resulting from chromatin preparation or fixation. Other developments based on cryo-focused ion beam (cryo-FIB) milling of intact, frozen cells¹¹⁸ or cryolysis¹¹⁹ also hold the potential of devising fixation-free versions of GAM and SPRITE that sample fractionated frozen nuclei.

Insights into chromatin organization

Each of the techniques available for studying 3D chromatin folding has provided important structural and functional insight into the different hierarchical levels of chromatin organization.

Chromosome territories and interchromosomal contacts.

FISH imaging shows that specific chromosomes occupy discrete non-random nuclear spaces during interphase, termed chromosome territories¹²⁰ (FIG. 1). Chromosome territories show cell type-dependent preferences in terms of both their radial position within the nucleus and their position relative to other chromosomes^{25,107,121}. Specific contacts can be detected at the interface between chromosome territories^{20,122,123}; overall, an estimated 20% of the volume of chromosome territories intermingles with other chromosome territories, often at their peripheries, both in human primary lymphocytes²⁴ and in *Drosophila melanogaster* cells^{25,108}. The extent of intermingling between chromosome territories directly correlates with translocation probabilities upon ionizing radiation damage, highlighting that the physical proximity between chromosomes affects their stability in response to DNA damage^{25,124,125}. The organization of chromosomes into discrete territories is also inferred from 3C-based and ligation-free approaches, as higher interaction frequencies are detected within chromosomes than between them (FIG. 1). 3C-based technologies have also detected contacts between chromosomes, and these have been successfully validated by imaging^{52,74,78,110,122,126–128}.

Chromatin hubs and compartments. The organization of chromosomes into subchromosomal domains has been extensively studied. For example, in mammalian cells, chromatin domains were observed in relation to replication origins, which contain many replicons and maintain their domain co-association across subsequent cell cycles¹²⁹. The compartmentalization of chromosomes into early and late replicating domains^{130–132} was also shown to be linked to transcriptional activity, with sites of active transcription occurring predominantly in early replicating domains¹³³. More recently, these observations have been largely confirmed by genome-wide assays to map replication and transcription, in which transcriptionally active and early replicating chromatin domains organize into separate subcompartments, distinct from late replicating domains^{134–136}. Early analyses of nuclear organization by electron and confocal microscopy had shown that chromatin occurs in highly condensed (heterochromatic) and less condensed (euchromatic) states¹³⁷, and revealed that transcription occurs in euchromatic areas of the nucleus¹³⁸. With the emergence of whole-genome 3C-based methodologies, such as Hi-C, the mapping of active and repressed chromatin states has become possible at the genome-wide scale, providing powerful insights into how gene expression relates to chromatin compaction. Application of principal component analysis to Hi-C data revealed a strong segregation of ligation events into two distinct compartments (A and B compartments) according to the activity state of the genomic regions⁴⁴. These compartments can also be seen in contact maps generated by ligation-free approaches^{10,11} (FIG. 1). Comparisons with linear maps of protein occupancy on chromatin helped reveal a strong relationship between the A compartment and transcriptionally active, open chromatin, as defined by DNase hypersensitivity, and the B compartment with closed chromatin, defined by repressive epigenetic marks of heterochromatin⁴⁴. Increased depth of Hi-C data sets has since allowed smaller subcompartments to be detected, which capture fine differences in replication timing as well as preferred associations with the nucleolus or the nuclear lamina⁶⁴.

Nuclear compartments or domains. Nuclear compartments are membrane-free organelles enriched for specific nuclear proteins and RNAs, which often have preferred associations with specific genomic regions and thereby influence the large-scale organization of chromosomes during interphase. They include the nucleolus, nuclear lamina, splicing speckles, paraspeckles, Cajal bodies, promyelocytic leukaemia bodies, Polycomb bodies, replication factories and transcription factories, which have all been described initially using microscopy^{139,140} (FIG. 1). For example, active ribosomal gene clusters are localized in the nucleolus, where the large ribosomal RNAs are transcribed, processed and assembled into pre-ribosomes¹⁴¹. Splicing speckles occupy internal nuclear positions, separate from the nuclear lamina and nucleoli, and bring together gene-dense regions^{91,142,143}. Association between specific genes at splicing speckles has been shown using imaging techniques, and has been confirmed at the genome-wide

level with SPRITE, which revealed that regions from different chromosomes come together at the same speckles¹¹. Genome-wide mapping of gene association with speckles has also recently been achieved by TSA-seq⁹². Fluorescence microscopy and electron microscopy showed that transcription itself occurs at discrete sites in the nucleus, termed transcription factories, which may organize active transcription units^{144–146}, with only a small proportion of transcriptional activity (~5–10%) being found immediately adjacent to the most prominent splicing speckles¹⁴⁶. Interestingly, the fraction of the genome that associates closely with slicing speckles has been shown by TSA-seq to contain highly transcribed genes and super-enhancers⁹², in keeping with previous imaging data¹⁴². Co-expressed genes can share the same transcription factory, which may be compatible with mechanisms of coordinated gene regulation via chromatin folding^{26,78,147,148}, but it remains unclear whether transcription factories are strictly specialized. Recent findings show that several factors involved in the transcription process, such as Pol II¹⁴⁹ or transcriptional co-activators BRD4 and MED1 (REF.¹⁵⁰), can form condensates by liquid–liquid phase separation, a process that may concentrate transcription factors and generate transcription factories. Moreover, the formation of nuclear condensates has been suggested as a general principle of nuclear body formation¹⁵¹. Clustering of distant genomic regions is not only mediated by transcription, but also occurs in the context of gene repression. Chromatin contacts at Polycomb bodies, which are repressive nuclear compartments, are a prominent example of gene clustering. In *D. melanogaster*, Polycomb-repressed Hox genes come together over a genomic distance of 10 Mb when they interact with a Polycomb body¹⁵². Other studies have reported long-range intrachromosomal and inter-chromosomal contacts between Polycomb-bound genes in human teratocarcinoma cells¹⁵³ and in mouse ESCs⁵².

Understanding how the preferential associations of genomic regions with specific nuclear domains relate to 3C-derived chromatin contacts remains a major challenge. Comparisons of genome-wide maps of lamina-associated domains⁹⁰ and Hi-C contacts show a strong coincidence between the transcriptionally inactive B compartment and the nuclear lamina^{64,94,154} or late replication domains¹³⁶. Repressive histone marks that define the heterochromatic B compartment are also strongly enriched at genomic regions that associate with the nucleolus¹⁵⁵, suggesting that the compacted B compartment is both situated at the nuclear periphery and clustered around the more central nucleoli, separated by the active, open A compartment. However, the bimodal separation of the A and B compartments derived from 3C-technologies should not be naively inferred as strictly active or silent chromatin. Genes can be activated in all areas of the nucleus, including at the nuclear lamina¹⁵⁶ or at centromeric regions¹⁵⁷, and gene positioning at the periphery does not always lead to gene inactivation^{158,159}. Heterochromatin domains also contain active sites of transcription¹⁶⁰. Consequently, a strict separation of the A and B compartments, as defined by 3C-approaches, seems unlikely, especially considering that contacts between compartments can be found in Hi-C maps⁹⁴ and

that they are found even more robustly using orthogonal methods, such as GAM, that do not rely on weak fixations¹⁰. These observations suggest that long-range gene-regulation mechanisms are complex, and not only depend on pairwise contacts between genomic regions but may also be influenced by the local nuclear environment where each region is located⁹⁵. The ongoing challenge of disentangling the direct functional relationships between the positions of genomic regions in the nucleus, their local and long-range contacts, and the state of gene activity is being addressed by analysing chromatin contacts at the single-cell level and with allele specificity, for example, using DNA-FISH²¹ or single-cell Hi-C^{73,74}.

TADs and loop domains. At smaller scales, chromosomes fold into self-associating chromatin domains, termed TADs^{24,94,161} (FIG. 1). Chromatin domains had been previously identified by microscopy but their detailed genomic composition was unclear. Since the discovery of TADs, the segmentation of the genome into megabase-sized domains has been extensively studied in several organisms and with different methodologies, leading to major breakthroughs in the discovery of mechanisms of disease caused by congenital genomic rearrangements^{3,47,162,163}. TADs often enclose clusters of co-regulated enhancers and promoters^{164,165}. Their size has been re-examined with the increasing resolution afforded by improved 3C-based assays, and found to vary from 40 kb to 3 Mb in the human genome⁶⁴, leading to the proposal of smaller loop domains as a sub-structure of TADs. Loop domains had been detected by microscopy before the emergence of 3C-technologies as DNA loops between transcriptionally active regions¹⁶⁶. Loop domains derived from 3C-based technologies often coincide with pairs of convergent CTCF binding sites, indicating that CTCF binding can contribute to the partition of specific regions of the genome into self-associating domains^{64,167–169}. Higher-order contacts between TADs have also been investigated, leading to the identification of metaTADs, which bring together distant TADs in cell type-specific patterns that relate to gene activity^{154,170}.

It has been debated whether TADs represent domains that exist predominantly across the cell population or represent an average of individual preferred contacts. Although interactions observed in single cells by single-cell Hi-C and imaging do not often identify whole TADs, the contacts detected frequently occur with the TAD coordinates defined by population Hi-C^{32,34,72}. However, this preference might not be as strong as anticipated. Imaging of chromatin contacts in mouse ESCs and oocytes showed that in 40% of cases 3D physical distances between regions that flank TAD borders are shorter than distances between regions within TADs⁷³, leading to highly variable contact clusters in individual cells that do not coincide with the positions of TADs in the cell population. This observation agrees with the detection of chromatin contacts between regions separated by TAD borders in single cells, often found at similar frequencies to regions within TADs²¹. However, it is particularly noteworthy that combining the single-cell Hi-C data results in the same TAD coordinates

observed in bulk population Hi-C, which supports the idea that TADs represent contact preferences of a cell population, rather than compact domains of chromatin in single cells^{73,171}.

Chromatin contacts between cis-regulatory elements.

Physical contacts between enhancers and promoters are essential for the transcription of genes⁸⁵ and can occur over distances ranging from less than 1 kb up to several megabases^{172–176} (FIG. 1). Genome-wide maps of candidate promoter–enhancer contacts can be created using high-resolution 3C-based methodologies that enrich for contacts mediated by Pol II or promoter histone marks, or that preferentially capture promoter-based contacts^{52,80,81}. Direct pairwise contacts between gene promoters and enhancers have become the most prominent concept of enhancer function, possibly as a result of the increased power of 3C-based technologies to detect local pairwise contacts rather than higher-order conformations. However, other mechanisms for regulating enhancer function are also emerging, which can involve formation of chromatin hubs, tethering of genes to active chromatin or nuclear environments^{156,158,177,178} and phase separation^{179,180}. An interesting study in budding yeast suggests homologue pairing as a mechanism for gene activation¹⁸¹. In the diploid yeast genome, upon glucose deprivation of the cell, both copies of the genomic locus containing the gene *TDA1* are relocalized to the nuclear periphery, where the homologues associate with each other and *TDA1* expression is activated. A more classical concept of gene regulation can be observed at developmental loci, where *cis*-regulatory contacts between enhancers and promoters are thought to occur most commonly within TADs^{48,162,182}. Although regulatory landscapes within TADs seem to be a common mechanism, genes themselves also contact each other across TAD boundaries over large genomic distances^{10,152,153,183}. Ligation-free methods, such as FISH, GAM and SPRITE,

all detect long-range contacts across TAD borders^{10,11,154}, and detailed analyses of Hi-C ligation frequencies also identify ligation events across TADs, over tens of megabases, that are statistically different from random contacts¹⁵⁴. The functional relevance of these contacts is a compelling question that is beginning to be addressed by developments that allow ectopic chromatin contacts to be engineered in the cell^{184,185}. The spatial and functional relationship between gene promoters that contact each other also remains poorly understood. Deletions of several gene promoters in the mouse ESC genome altered the expression of nearby genes¹⁸⁶. This observation suggests that genes themselves may act as enhancers for other genes, possibly by recruiting *cis*-regulatory signals, and supports the concept that clustering of genes in transcription factories has regulatory functions.

Conclusions

The development of genome-wide approaches for studying 3D genome folding have revolutionized our ability to understand the regulatory content of the linear genomic sequence. Alongside 3C-based methods, the recent development of orthogonal technologies to map chromatin contacts brings us closer to uncovering 3D genome folding architectures with unprecedented detail, at all genomic scales and with single-cell resolution. The ongoing revolution in live-cell imaging, including improvements to the number of genomic loci that can be tagged simultaneously, will provide a deeper mechanistic understanding of how 3D folding structures are formed and disassembled, and how they contribute to genome stability gene expression, and of homeostatic changes in cell states in response to stimuli. Ultimately, the ability to detect changes in chromosome topology will open new avenues for disease diagnostics, disease target discovery and many other applications.

Published online: 17 December 2019

- Pombo, A. & Dillon, N. Three-dimensional genome architecture: players and mechanisms. *Nat. Rev. Mol. Cell Biol.* **16**, 245–257 (2015).
- Dekker, J. et al. The 4D nucleome project. *Nature* **549**, 219–226 (2017).
- Spielmann, M., Lupiáñez, D. G. & Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018).
This review explains the influence of disrupted chromatin folding on gene regulation in disease with examples from developmental disorders.
- Krijger, P. H. & de Laat, W. Regulation of disease-associated gene expression in the 3D genome. *Nat. Rev. Mol. Cell Biol.* **17**, 771–782 (2016).
- Gall, J. G. & Pardue, M. L. Formation and detection of RNA–DNA hybrid molecules in cytological preparations. *Proc. Natl Acad. Sci. USA* **63**, 378–383 (1969).
- Speicher, M. R., Gwyn Ballard, S. & Ward, D. C. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat. Genet.* **12**, 368–375 (1996).
- Wang, S. et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016).
- Ma, H., Reyes-Gutiérrez, P. & Pederson, T. Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. *Proc. Natl Acad. Sci. USA* **110**, 21048–21053 (2013).
- Ma, H. et al. Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.* **34**, 528–530 (2016).
- Beagrie, R. A. et al. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519–524 (2017).
This study introduces GAM, a ligation-free technique to map chromatin contacts genome-wide. GAM confirms the presence of TADs and reveals multiway contacts between super-enhancers spanning tens of megabases.
- Quinodoz, S. A. et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757.e24 (2018).
This paper describes SPRITE, a ligation-free approach to map chromatin interactions and DNA–RNA contacts across the entire genome. SPRITE detects chromosomal hubs at nuclear bodies that bring together regions from different chromosomes.
- Zheng, M. et al. Multiplex chromatin interactions with single-molecule precision. *Nature* **566**, 558–562 (2019).
- Cutter, A. R. & Hayes, J. J. A brief review of nucleosome structure. *FEBS Lett.* **589**, 2914–2922 (2015).
- Robinett, C. C. et al. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J. Cell Biol.* **135**, 1685–1700 (1996).
- Belmont, A. S. & Straight, A. F. In vivo visualization of chromosomes using lac operator-repressor binding. *Trends Cell Biol.* **8**, 121–124 (1998).
- Lucas, J. S., Zhang, Y., Dudko, O. K. & Murre, C. 3D trajectories adopted by coding and regulatory DNA elements: first-passage times for genomic interactions. *Cell* **158**, 339–352 (2014).
- Germier, T., Sylvain, A., Silvia, K., David, L. & Kerstin, B. Real-time imaging of specific genomic loci in eukaryotic cells using the ANCHOR DNA labelling system. *Methods* **142**, 16–23 (2018).
- Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L. & Rinn, J. L. A TAD boundary is preserved upon deletion of the CTCF-rich *Firre* locus. *Nat. Commun.* **9**, 1444–1455 (2018).
- Barbieri, M. et al. Active and poised promoter states drive folding of the extended *HoxB* locus in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **24**, 515–524 (2017).
- Maass, P. G., Barutcu, A. R., Weiner, C. L. & Rinn, J. L. Inter-chromosomal contact properties in live-cell imaging and in Hi-C. *Mol. Cell* **69**, 1039–1045.e3 (2018).
- Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019).
- Boettiger, A. N. et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
- Luperchio, T. R. et al. Chromosome conformation paints reveal the role of lamina association in genome organization and regulation. Preprint at *bioRxiv* <https://doi.org/10.1101/122226> (2017).
- Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the *X*-inactivation centre. *Nature* **485**, 381–385 (2012).
This study uses 5C to identify the compartmentalization of chromosomes into TADs.
- Branco, M. R. & Pombo, A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent

- associations. *PLoS Biol.* **4**, e1380780–e1380788 (2006).
- This microscopy study of contacts between chromosomes uses cryo-FISH to show the previously underestimated extent of chromosome intermingling between chromosome territories.**
26. Ferrai, C. et al. Poised transcription factories prime silent uPA gene prior to activation. *PLoS Biol.* **8**, e1000270 (2010).
 27. Simonis, M. et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nat. Genet.* **38**, 1348–1354 (2006).
 28. Beliveau, B. J. et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl Acad. Sci. USA* **109**, 21301–21306 (2012).
 29. Boyle, S., Rodesch, M. J., Halvensleben, H. A., Jeddeloh, J. A. & Bickmore, W. A. Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. *Chromosome Res.* **19**, 901–909 (2011).
 30. Beliveau, B. J. et al. Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes. *Nat. Commun.* **6**, 7147–7160 (2015).
 31. Gnirke, A. et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* **27**, 182–189 (2009).
 32. Bintu, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018).
 33. Ni, Y. et al. Super-resolution imaging of a 2.5 kb non-repetitive DNA in situ in the nuclear genome using molecular beacon probes. *eLife* **6**, 21660 (2017).
 - This elegant study uses super-resolution microscopy and DNA-FISH to fine-map enhancer–promoter contacts.**
 34. Stevens, T. J. et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59–64 (2017).
 35. Gibcus, J. H. et al. A pathway for mitotic chromosome formation. *Science* **359**, eaao6135 (2018).
 36. Chen, B. et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
 37. Shao, S. et al. Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system. *Nucleic Acids Res.* **44**, e86–e98 (2016).
 38. Fu, Y. et al. CRISPR–dCas9 and sgRNA scaffolds enable dual-colour live imaging of satellite sequences and repeat-enriched individual loci. *Nat. Commun.* **7**, 11707–11714 (2016).
 39. Wang, S., Su, J. H., Zhang, F. & Zhuang, X. An RNA-aptamer-based two-color CRISPR labeling system. *Sci. Rep.* **6**, 26857–26863 (2016).
 40. Gu, B. et al. Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* **359**, 1050–1055 (2018).
 41. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
 42. Zhao, Z. et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* **38**, 1341–1347 (2006).
 43. Dostie, J. et al. Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–1309 (2006).
 44. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
 - This article describes Hi-C, a 3C-based approach to map chromatin contacts genome-wide using selection of ligated DNA fragments.**
 45. Denker, A. & de Laat, W. The second decade of 3C technologies: detailed insights into nuclear organization. *Genes Dev.* **30**, 1357–1382 (2016).
 46. van de Werken, H. J. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969–972 (2012).
 47. Franke, M. et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265–269 (2016).
 48. Symmons, O. et al. The Shh topological domain facilitates the action of remote enhancers by reducing the effects of genomic distances. *Dev. Cell* **39**, 529–543 (2016).
 49. Loviglio, M. N. et al. Chromosomal contacts connect loci associated with autism, BMI and head circumference phenotypes. *Mol. Psychiatry* **22**, 836–849 (2017).
 50. Kundu, S. et al. Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. *Mol. Cell* **65**, 432–446.e5 (2017).
 51. Hughes, J. R. et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205–212 (2014).
 52. Mifsud, B. et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.* **47**, 598–606 (2015).
 53. Liu, X. et al. In situ capture of chromatin interactions by biotinylated dCas9. *Cell* **170**, 1028–1043.e19 (2017).
 54. Andrey, G. et al. Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding. *Genome Res.* **27**, 223–233 (2017).
 55. Schoenfelder, S., Javierre, B. M., Furlan-Magaril, M., Wingett, S. W. & Fraser, P. Promoter capture Hi-C: high-resolution, genome-wide profiling of promoter interactions. *J. Vis. Exp.* **136**, e57320 (2018).
 56. Belton, J. M. et al. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* **58**, 268–276 (2012).
 57. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F. & Chen, L. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.* **30**, 90–98 (2011).
 58. Rodley, C. D. M., Bertels, F., Jones, B. & O’Sullivan, J. M. Global identification of yeast chromosome interactions using genome conformation capture. *Fungal Genet. Biol.* **46**, 879–886 (2009).
 - This paper describes GCC, a 3C-based approach to map chromatin contacts genome-wide without selection of ligated DNA fragments.**
 59. Servant, N., Varoquaux, N., Heard, E., Barillot, E. & Vert, J.-P. Effective normalization for copy number variation in Hi-C data. *BMC Bioinformatics* **19**, 313–313 (2018).
 60. Dixon, J. R. et al. Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* **50**, 1388–1398 (2018).
 61. Vidal, E. et al. OneD: increasing reproducibility of Hi-C samples with abnormal karyotypes. *Nucleic Acids Res.* **46**, e49–e58 (2018).
 62. Ma, W. et al. Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. *Nat. Methods* **12**, 71–78 (2015).
 63. Ma, W. et al. Using DNase Hi-C techniques to map global and local three-dimensional genome architecture at high resolution. *Methods* **142**, 59–73 (2018).
 64. Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
 - This article describes a high-resolution contact map of the human genome, which reveals the presence of loop domains and introduces the concept of loop formation between divergent CTCF sites.**
 65. Hsieh, T.-H. S. et al. Mapping nucleosome resolution chromosome folding in yeast by micro-C. *Cell* **162**, 108–119 (2015).
 66. Hsieh, T. S., Fudenberg, G., Goloborodko, A. & Rando, O. J. Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome. *Nat. Methods* **13**, 1009–1011 (2016).
 67. Hsieh, T.-H. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. Preprint at *bioRxiv* <https://doi.org/10.1101/638775> (2019).
 68. Li, T., Jia, L., Cao, Y., Chen, Q. & Li, C. OCEAN-C: mapping hubs of open chromatin interactions across the genome reveals gene regulatory networks. *Genome Biol.* **19**, 54–68 (2018).
 69. Gavrillo, A. A. et al. Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. *Nucleic Acids Res.* **41**, 3563–3575 (2013).
 - This paper presents electron microscopy and confocal microscopy images showing the changes that occur in chromatin ultrastructure during a proximity-ligation assay.**
 70. Lu, L., Liu, X., Peng, J., Li, Y. & Jin, F. Easy Hi-C: a simple efficient protocol for 3D genome mapping in small cell populations. Preprint at *bioRxiv* <https://doi.org/10.1101/245688> (2018).
 71. Nagano, T. et al. Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. *Nat. Protoc.* **10**, 1986–2003 (2015).
 72. Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013).
 73. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
 - This study uses single-cell Hi-C and DNA-FISH to study differences in 3D genome folding between oocytes and zygotes, and reveals the absence of chromatin compartments in the oocyte, as well as single-cell heterogeneity in TAD organization.**
 74. Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
 - This paper describes the use of single-cell Hi-C to map chromatin contacts throughout the cell cycle in thousands of individual cells, which shows the emergence of chromatin structures after mitosis and the temporal dynamics of different chromatin topologies.**
 75. Ramani, V. et al. Massively multiplex single-cell Hi-C. *Nat. Methods* **14**, 263–266 (2017).
 76. Tan, L., Xing, D., Chang, C.-H., Li, H. & Xie, X. S. Three-dimensional genome structures of single diploid human cells. *Science* **361**, 924–928 (2018).
 77. Horike, S., Cai, S., Miyano, M., Cheng, J. F. & Kohwi-Shigematsu, T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.* **37**, 31–40 (2005).
 78. Schoenfelder, S. et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* **42**, 53–61 (2010).
 79. Fullwood, M. J. et al. An oestrogen-receptor- α -bound human chromatin interactome. *Nature* **462**, 58–64 (2009).
 80. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
 81. Fang, R. et al. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.* **26**, 1345–1348 (2016).
 82. Davies, J. O., Oudelaar, A. M., Higgs, D. R. & Hughes, J. R. How best to identify chromosomal interactions: a comparison of approaches. *Nat. Methods* **14**, 125–134 (2017).
 83. Zhang, Y. et al. Enhancing Hi-C data resolution with deep convolutional neural network HiCPlus. *Nat. Commun.* **9**, 750–758 (2018).
 84. O’Sullivan, J. M., Hendy, M. D., Pichugin, T., Wake, G. C. & Langowski, J. The statistical-mechanics of chromosome conformation capture. *Nucleus* **4**, 390–398 (2013).
 85. Chen, H. et al. Dynamic interplay between enhancer–promoter topology and gene activity. *Nat. Genet.* **50**, 1296–1303 (2018).
 - In this study, live-cell imaging shows that transcription directly depends on contact between enhancers and promoters.**
 86. van Steensel, B. & Henikoff, S. Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. *Nat. Biotechnol.* **18**, 424–428 (2000).
 87. Vogel, M. J., Peric-Hupkes, D. & van Steensel, B. Detection of in vivo protein–DNA interactions using DamID in mammalian cells. *Nat. Protoc.* **2**, 1467–1478 (2007).
 88. Marshall, O. J., Southall, T. D., Cheatham, S. W. & Brand, A. H. Cell-type-specific profiling of protein–DNA interactions without cell isolation using targeted DamID with next-generation sequencing. *Nat. Protoc.* **11**, 1586–1598 (2016).
 89. Peric-Hupkes, D. et al. Molecular maps of the reorganization of genome–nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010).
 90. Guelen, L. et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948–951 (2008).
 - This article describes DamID, an elegant approach to map chromatin contacts at the nuclear lamina.**
 91. Spector, D. L. & Lamond, A. I. Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* **3**, a000646 (2011).
 92. Chen, Y. et al. Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *J. Cell Biol.* **270**, 4025–4048 (2018).
 - This article introduces TSA-seq, a genome-wide sequencing technique, to map spatial distances between DNA and nuclear bodies, such as splicing speckles.**

93. Redolfi, J. et al. DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat. Struct. Mol. Biol.* **26**, 471–480 (2019). **This article introduces DamC, an approach that allows mapping of chromatin contacts in vivo and without crosslinking and ligation.**
94. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012). **In this article, Hi-C experiments reveal the compartmentalization of chromosomes into TADs.**
95. Pombo, A. & Branco, M. R. Functional organisation of the genome during interphase. *Curr. Opin. Genet. Dev.* **17**, 451–455 (2007).
96. Li, Y. et al. The effects of chemical fixation on the cellular nanostructure. *Exp. Cell Res.* **358**, 253–259 (2017).
97. Oudelaar, A. M., Davies, J. O. J., Downes, D. J., Higgs, D. R. & Hughes, J. R. Robust detection of chromosomal interactions from small numbers of cells using low-input Capture-C. *Nucleic Acids Res.* **45**, e184–e192 (2017).
98. Guillot, P. V., Xie, S. Q., Hollinshead, M. & Pombo, A. Fixation-induced redistribution of hyperphosphorylated RNA polymerase II in the nucleus of human cells. *Exp. Cell Res.* **295**, 460–468 (2004).
99. Solovei, I. et al. Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence in situ hybridization (3D-FISH). *Exp. Cell Res.* **276**, 10–23 (2002).
100. Markaki, Y. et al. The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture: 3D structured illumination microscopy of defined chromosomal structures visualized by 3D (immuno)-FISH opens new perspectives for studies of nuclear architecture. *Bioessays* **34**, 412–426 (2012).
101. Xie, S. Q., Lavitas, L. M. & Pombo, A. CryoFISH: fluorescence in situ hybridization on ultrathin cryosections. *Methods Mol. Biol.* **659**, 219–230 (2010).
102. Brown, J. M. et al. A tissue-specific self-interacting chromatin domain forms independently of enhancer–promoter interactions. *Nat. Commun.* **9**, 3849–3863 (2018).
103. Allahyar, A. et al. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat. Genet.* **50**, 1151–1160 (2018).
104. Olivares-Chauvet, P. et al. Capturing pairwise and multi-way chromosomal conformations using chromosomal walks. *Nature* **540**, 296–300 (2016).
105. Darrow, E. M. et al. Deletion of DXZ4 on the human inactive X chromosome alters higher-order genome architecture. *Proc. Natl Acad. Sci. USA* **113**, E4504–E4512 (2016).
106. Oudelaar, A. M. et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat. Genet.* **50**, 1744–1751 (2018).
107. Branco, M. R., Branco, T., Ramirez, F. & Pombo, A. Changes in chromosome organization during PHA-activation of resting human lymphocytes measured by cryo-FISH. *Chromosome Res.* **16**, 413–426 (2008).
108. Rosin, L. F., Nguyen, S. C. & Joyce, E. F. Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in *Drosophila* nuclei. *PLoS Genet.* **14**, e1007393 (2018).
109. Loviglio, M. N. et al. Chromosomal contacts connect loci associated with autism, BMI and head circumference phenotypes. *Mol. Psychiatry* **22**, 836–849 (2016).
110. Spiliarakis, C. G., Lalioti, M. D., Town, T., Lee, G. R. & Flavell, R. A. Interchromosomal associations between alternatively expressed loci. *Nature* **435**, 637–645 (2005).
111. Monahan, K., Horta, A. & Lomvardas, S. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* **565**, 448–453 (2019).
112. Hakim, O. et al. Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. *Genome Res.* **21**, 697–706 (2011).
113. Giorgetti, L. & Heard, E. Closing the loop: 3C versus DNA FISH. *Genome Biol.* **17**, 215–223 (2016).
114. Tang, Z. et al. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* **163**, 1611–1627 (2015).
115. Wang, X. T., Dong, P. F., Zhang, H. Y. & Peng, C. Structural heterogeneity and functional diversity of topologically associating domains in mammalian genomes. *Nucleic Acids Res.* **43**, 7237–7246 (2015). **This paper describes a chromosome-wide effort to map TADs using DNA-FISH. In addition to revealing the single-cell behaviour of TADs, the paper shows that the imaging data have a high correlation with Hi-C data.**
116. Fudenberg, G. & Imakaev, M. FISH-ing for captured contacts: towards reconciling FISH and 3C. *Nat. Methods* **14**, 673–678 (2017).
117. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320.e24 (2017).
118. Mahamid, J. et al. Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science* **351**, 969–972 (2016).
119. Aitchison, J. D. & Rout, M. P. The interactome challenge. *J. Cell Biol.* **211**, 729 (2015).
120. Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2**, a003889 (2010).
121. Parada, L. & Misteli, T. Chromosome positioning in the interphase nucleus. *Trends Cell Biol.* **12**, 425–432 (2002).
122. Hacısuleyman, E. et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* **21**, 198–206 (2014).
123. Maass, P. G. et al. Reorganization of inter-chromosomal interactions in the 2q37-deletion syndrome. *EMBO J.* **37**, e96257 (2018).
124. Maharana, S. et al. Chromosome intermingling — the physical basis of chromosome organization in differentiated cells. *Nucleic Acids Res.* **44**, 5148–5160 (2016).
125. Zhang, Y. et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* **148**, 908–921 (2012).
126. Lomvardas, S. et al. Interchromosomal interactions and olfactory receptor choice. *Cell* **126**, 403–413 (2006).
127. Cairns, J. et al. CHICAGO: robust detection of DNA looping interactions in capture Hi-C data. *Genome Biol.* **17**, 127–143 (2016).
128. Fanucchi, S., Shibayama, Y., Burd, S., Weinberg, M. S. & Mhlanga, M. M. Chromosomal contact permits transcription between coregulated genes. *Cell* **155**, 606–620 (2013).
129. Jackson, D. A. & Pombo, A. Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J. Cell Biol.* **140**, 1285–1295 (1998).
130. Zink, D., Bornfleth, H., Visser, A., Cremer, C. & Cremer, T. Organization of early and late replicating DNA in human chromosome territories. *Exp. Cell Res.* **247**, 176–188 (1999).
131. Visser, A. E. et al. Spatial distributions of early and late replicating chromatin in interphase chromosome territories. *Exp. Cell Res.* **245**, 398–407 (1998).
132. Ferreira, J., Paoletta, G., Ramos, C. & Lamond, A. I. Spatial organization of large-scale chromatin domains in the nucleus: a magnified view of single chromosome territories. *J. Cell Biol.* **139**, 1597–1610 (1997).
133. Sadoni, N. et al. Nuclear organization of mammalian genomes. *J. Cell Biol.* **146**, 1211–1226 (1999).
134. Hiratani, I. et al. Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol.* **6**, e245 (2008).
135. Schwaiger, M. et al. Chromatin state marks cell-type- and gender-specific replication of the *Drosophila* genome. *Genes Dev.* **23**, 589–601 (2009).
136. Pope, B. D. et al. Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402–405 (2014).
137. Monneron, A. & Bernhard, W. Fine structural organization of the interphase nucleus in some mammalian cells. *J. Ultrastruct. Res.* **27**, 266–288 (1969).
138. Verschure, P. J., van der Kraan, I., Manders, E. M. M. & van Driel, R. Spatial relationship between transcription sites and chromosome territories. *J. Cell Biol.* **147**, 13–24 (1999).
139. Dunder, M. & Misteli, T. Biogenesis of nuclear bodies. *Cold Spring Harb. Perspect. Biol.* **2**, a000711 (2010).
140. Mao, Y. S., Zhang, B. & Spector, D. L. Biogenesis and function of nuclear bodies. *Trends Genet.* **27**, 295–306 (2011).
141. Pederson, T. The nucleolus. *Cold Spring Harb. Perspect. Biol.* **3**, a000638 (2011).
142. Shopland, L. S., Johnson, C. V., Byron, M., McNeil, J. & Lawrence, J. B. Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. *J. Cell Biol.* **162**, 981–990 (2003).
143. Brown, J. M. et al. Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J. Cell Biol.* **182**, 1083–1097 (2008).
144. Iborra, F. J., Pombo, A., Jackson, D. A. & Cook, P. R. Active RNA polymerases are localized within discrete transcription “factories” in human nuclei. *J. Cell Sci.* **109**, 1427–1436 (1996).
145. Pombo, A. et al. Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J.* **18**, 2241–2253 (1999).
146. Xie, S. Q., Martin, S., Guillot, P. V., Bentley, D. L. & Pombo, A. Splicing speckles are not reservoirs of RNA polymerase II, but contain an inactive form, phosphorylated on serine2 residues of the C-terminal domain. *Mol. Biol. Cell* **17**, 1723–1733 (2006).
147. Osborne, C. S. et al. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* **36**, 1065–1071 (2004).
148. Osborne, C. S. & Eskiw, C. H. Where shall we meet? A role for genome organization and nuclear sub-compartments in mediating interchromosomal interactions. *J. Cell Biochem.* **104**, 1553–1561 (2008).
149. Boehning, M. et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).
150. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, aar3958 (2018).
151. Banani, S. F. et al. Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 (2016).
152. Bantignies, F. et al. Polcomb-dependent regulatory contacts between distant Hox loci in *Drosophila*. *Cell* **144**, 214–226 (2011).
153. Tiwari, V. K., Cope, L., McGarvey, K. M., Ohm, J. E. & Baylin, S. B. A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. *Genome Res.* **18**, 1171–1179 (2008).
154. Fraser, J. et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol. Syst. Biol.* **11**, 852–865 (2015).
155. Nemeth, A. et al. Initial genomics of the human nucleolus. *PLoS Genet.* **6**, e1000889 (2010).
156. Kumaran, R. I. & Spector, D. L. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J. Cell Biol.* **180**, 51–65 (2008).
157. Sabbattini, P., Georgiou, A., Sinclair, C. & Dillon, N. Analysis of mice with single and multiple copies of transgenes reveals a novel arrangement for the $\lambda 5$ -VpreB1 locus control region. *Mol. Cell Biol.* **19**, 671–679 (1999).
158. Finlan, L. E. et al. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet.* **4**, e1000039 (2008). **This article demonstrates that the nuclear environment influences gene expression; bringing genes into the repressive context of the nuclear lamina can lead to downregulation of their expression.**
159. Wang, H. et al. CRISPR-mediated programmable 3D genome positioning and nuclear organization. *Cell* **175**, 1405–1417.e14 (2018).
160. Grewal, S. I. & Elgin, S. C. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**, 399–406 (2007).
161. Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
162. Lupianez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene–enhancer interactions. *Cell* **161**, 1012–1025 (2015).
163. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
164. Shen, Y. et al. A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
165. Symmons, O. et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **24**, 390–400 (2014).
166. Jackson, D. A., Bartlett, J. & Cook, P. R. Sequences attaching loops of nuclear and mitochondrial DNA to underlying structures in human cells: the role of transcription units. *Nucleic Acids Res.* **24**, 1212–1219 (1996).
167. de Wit, E. et al. CTCF binding polarity determines chromatin looping. *Mol. Cell* **60**, 676–684 (2015).
168. Gomez-Marin, C. et al. Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *Proc. Natl Acad. Sci. USA* **112**, 7542–7547 (2015).
169. Vietri Rudan, M. et al. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* **10**, 1297–1309 (2015).

170. Weinreb, C. & Raphael, B. J. Identification of hierarchical chromatin domains. *Bioinformatics* **32**, 1601–1609 (2016).
171. Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
172. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active β -globin locus. *Mol. Cell* **10**, 1453–1465 (2002).
173. Lettice, L. A. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735 (2003).
174. Nobrega, M. A., Ovcharenko, I., Afzal, V. & Rubin, E. M. Scanning human gene deserts for long-range enhancers. *Science* **302**, 413–413 (2003).
175. Qin, Y. et al. Long-range activation of Sox9 in Odd Sex (Ods) mice. *Hum. Mol. Genet.* **13**, 1213–1218 (2004).
176. Javierre, B. M. et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* **167**, 1369–1384.e19 (2016).
177. Zullo, J. M. et al. DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. *Cell* **149**, 1474–1487 (2012).
178. Reddy, K. L., Zullo, J. M., Bertolino, E. & Singh, H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**, 243–247 (2008).
179. Nott, T. J. et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947 (2015).
180. Strom, A. R. et al. Phase separation drives heterochromatin domain formation. *Nature* **547**, 241–245 (2017).
This article introduces phase separation as a model for heterochromatin formation in the nucleus by showing liquid–liquid phase separation of the heterochromatic protein HP1 in vitro.
181. Kim, S. et al. The dynamic three-dimensional organization of the diploid yeast genome. *eLife* **6**, 23623–23645 (2017).
182. Chetverina, D., Aoki, T., Erokhin, M., Georgiev, P. & Schedl, P. Making connections: insulators organize eukaryotic chromosomes into independent cis-regulatory networks. *Bioessays* **36**, 163–172 (2014).
183. Schoenfelder, S. et al. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* **25**, 582–597 (2015).
184. Deng, W. et al. Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849–860 (2014).
185. Kim, J. H. et al. LADL: light-activated dynamic looping for endogenous gene expression control. *Nat. Methods* **16**, 633–639 (2019).
186. Engreitz, J. M. et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).
This study shows that regulation of genes can occur through the promoters of nearby genes, highlighting the functional importance of transcription factories.
187. Mayer, R. et al. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol.* **6**, 44–65 (2005).
188. Ahmed, K. et al. Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLOS ONE* **5**, e10531 (2010).
189. Li, Y. et al. CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells. *PLOS ONE* **9**, e114485 (2014).
190. Kerpedjiev, P. et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol.* **19**, 125–136 (2018).
191. Bonev, B. et al. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572.e24 (2017).
192. Naumova, N., Smith, E. M., Zhan, Y. & Dekker, J. Analysis of long-range chromatin interactions using chromosome conformation capture. *Methods* **58**, 192–203 (2012).
193. van de Werken, H. J. et al. 4C technology: protocols and data analysis. *Methods Enzymol.* **513**, 89–112 (2012).
194. Schwartzman, O. et al. UMI-4C for quantitative and targeted chromosomal contact profiling. *Nat. Methods* **13**, 685–691 (2016).
195. Dostie, J. & Dekker, J. Mapping networks of physical interactions between genomic elements using 5C technology. *Nat. Protoc.* **2**, 988–1002 (2007).
196. Kim, J. H. et al. 5C-ID: increased resolution chromosome-conformation-capture-carbon-copy with in situ 3C and double alternating primer design. *Methods* **142**, 39–46 (2018).
197. Belaghal, H., Dekker, J. & Gibcus, J. H. Hi-C 2.0: an optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56–65 (2017).
198. Li, X. et al. Long-read ChIA-PET for base-pair-resolution mapping of haplotype-specific chromatin interactions. *Nat. Protoc.* **12**, 899–915 (2017).
199. Davies, J. O. et al. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat. Methods* **13**, 74–80 (2016).
200. Croft, J. A. et al. Differences in the localization and morphology of chromosomes in the human nucleus. *J. Cell Biol.* **145**, 1119–1131 (1999).
201. Cremer, M. et al. Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. *Methods Mol. Biol.* **463**, 205–239 (2008).
202. Chen, B. et al. Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res.* **44**, e75–e87 (2016).
203. Takei, Y., Shah, S., Harvey, S., Qi, L. S. & Cai, L. Multiplexed dynamic imaging of genomic loci by combined CRISPR imaging and DNA sequential FISH. *Biophys. J.* **112**, 1773–1776 (2017).
204. Shimizu, N., Maekawa, M., Asai, S. & Shimizu, Y. Multicolor FISHs for simultaneous detection of genes and DNA segments on human chromosomes. *Chromosome Res.* **23**, 649–662 (2015).
205. Müller, S., Neusser, M. & Wienberg, J. Towards unlimited colors for fluorescence in-situ hybridization (FISH). *Chromosome Res.* **10**, 223–232 (2002).
206. Hepperger, C., Otten, S., von Hase, J. & Dietzel, S. Preservation of large-scale chromatin structure in FISH experiments. *Chromosoma* **116**, 117–133 (2007).
207. Volpi, E. V. 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. *J. Cell Sci.* **113**, 1565–1576 (2000).
208. Williams, R. R., Broad, S., Sheer, D. & Ragoussis, J. Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. *Exp. Cell Res.* **272**, 163–175 (2002).
209. Chambeyron, S. & Bickmore, W. A. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev.* **18**, 1119–1130 (2004).
210. Raap, A. K., Marijnen, J. G., Vrolijk, J. & van der Ploeg, M. Denaturation, renaturation, and loss of DNA during in situ hybridization procedures. *Cytometry* **7**, 235–242 (1986).

Acknowledgements

The authors thank the Helmholtz Association (Germany) for support, C. Thieme (our laboratory) for help plotting GAM and SPRITE contact maps in Figs 1 and 5a,b, and S. Quinodoz (M. Guttman laboratory) for sharing SPRITE contact cluster data (Figs 1 and 5b). A.P. acknowledges support from the National Institutes of Health Common Fund 4D Nucleome Program grant U54DK107977. The authors apologize to the many scientists whose studies were not discussed in our review due to length constraints.

Author contributions

All authors contributed to all aspects of the article.

Competing interests

A.P. has filed a patent application on GAM: Pombo, A., Edwards, P. A. W., Nicodemi, M., Beagrie, R. A. & Scialdone, A. Patent application on 'Genome Architecture Mapping'. Patent PCT/EP2015/079413 (2015).

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2019