



HAL
open science

Understanding 3D genome organization by multidisciplinary methods

Ivana Jerković, Giacomo Cavalli

► **To cite this version:**

Ivana Jerković, Giacomo Cavalli. Understanding 3D genome organization by multidisciplinary methods. Nature Reviews Molecular Cell Biology, 2021, 10.1038/s41580-021-00362-w . hal-03269159

HAL Id: hal-03269159

<https://hal.science/hal-03269159>

Submitted on 23 Jun 2021

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

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

1 Understanding 3D genome organization by multidisciplinary methods

2 Ivana Jerković and Giacomo Cavalli

3
4 Institute of Human Genetics, CNRS, University of Montpellier, Montpellier, France

5 giacomo.cavalli@igh.cnrs.fr

6 7 8 **Abstract**

9 Understanding how chromatin is folded in the nucleus is fundamental to understanding its function. Although 3D
10 genome organization has been historically difficult to study owing to lack of relevant methodologies, major
11 technological breakthroughs in genome-wide mapping of chromatin contacts and advances in imaging technologies
12 in the 21st century considerably improved our understanding of chromosome conformation and nuclear architecture.
13 In this Review, we discuss methods of 3D genome organization analysis, including **sequencing-based techniques**,
14 such as Hi-C and its derivatives, micro-C, DamID and others; **microscopy-based techniques**, such as super-resolution
15 imaging coupled with fluorescent *in situ* hybridization (FISH), multiplex FISH, *in situ* genome sequencing and live
16 microscopy methods; and **computational and modeling approaches**. We describe the most commonly used
17 techniques and their contribution to our current knowledge of nuclear architecture and, finally, we provide a
18 perspective on up-and-coming methods that open possibilities for future major discoveries.

19 20 Introduction

21 Euchromatin, heterochromatin and the hypothesis of individual-chromosome territories in the nucleus have been
22 observed and suggested already at the end of the 19th and the beginning of the 20th century using light microscopy
23 and chromatin dyes ¹⁻³. However, by the middle of 20th century, the chromosome territories hypothesis was largely
24 abandoned as conventional electron microscopy failed to confirm it. It was not until the 1980s and the development
25 of a novel imaging technique named fluorescence in-situ hybridization (FISH), that the chromosome territories
26 hypothesis was finally validated, thereby instigating the study of nuclear architecture as we know it today ⁴⁻⁶. The
27 FISH method demonstrated the existence of chromosome territories and chromosome intermingling at territory
28 edges, and also indicated that chromosomal regions rich in active genes largely reside at the nuclear interior,
29 whereas chromosomal regions rich in inactive genes largely reside at the nuclear periphery ⁷⁻¹⁶.

30 About two decades later, completion of sequencing of the human genome and the subsequent genome-wide
31 characterization of genetic variations and epigenetic transcription regulation propelled the study of nuclear
32 architecture into a new era. One breakthrough was the development of new techniques based on the principles of
33 ligation of linearly-distal genomic regions that come into 3D spatial proximity in order to map genome organization
34 and, at the same time, assay the functionality of this 3D organization ^{17,18}. However, despite the tremendous
35 advancements made in this relative short time, the full complexity of the biophysical principles underlying the strong

36 3D compaction of the long linear genomic DNA into a micrometer-sized nucleus remain to be elucidated.
37 Furthermore, we do not know the rules by which the structures imposed by general biophysical laws can be modified
38 at specific genes in order to drive changes in gene expression programs that underlie cell fate and plasticity. These
39 major challenges stimulate methodological improvements and invention of new experimental techniques and
40 modeling approaches. On the other hand, fast and numerous technical developments generate confusion in the field
41 of genome organization research, since it is unclear what information can each method provide, and the scarce
42 comparison of concurrent methods hampers progress. Given the technological challenges and limitations of each of
43 technology, the multiscale nature of genome organization and the multifaceted regulation of genome function, only
44 the deployment of the full discourse of experimental and theoretical genome analysis approaches will allow us to
45 reach a complete appreciation of genome function and the capacity to harness it in order to deliver not only
46 fundamental knowledge but also valuable biomedical applications.
47 In this Review, we discuss techniques for high-throughput chromatin contacts analysis and highly multiplexed, super-
48 resolution and live-imaging methods. The applicability of these techniques is intimately linked with computational
49 tools, including machine learning and mathematical modeling based on **first principles** or driven by quantitative data.
50 We discuss relevant technological progress, provide a view of the current state of 3D genome organization research,
51 and discuss promising future developments.

52 **Studying the multilayered 3D genome**

53 The idea of the nucleus as a highly organized organelle existed for over a century¹⁻³. However, so far, we grasp only
54 a part of the principles that govern nuclear organization, and the emergence of new evidence is tightly connected
55 with the development of new methods.

56 A major breakthrough in chromatin biology was the establishment of chromosome conformation capture (3C) — a
57 nuclear ligation assay in conjunction with PCR, which marked the beginning of the era of high-throughput next-
58 generation sequencing-based techniques for the investigation of chromosome conformation^{17,18}. Indeed, a series
59 of 3C derivatives (from here on defined as C-based techniques) were developed to assay contact frequency between
60 multiple genomic loci, including circular 3C (4C)¹⁹, which measures interaction frequencies of one locus with many
61 loci ('one-vs-many'); many-vs-many assays (3C carbon copy (5C)²⁰, Capture-C²¹⁻²⁵, Capture-Hi-C²⁶, etc.) and
62 genome-wide, all-vs-all assays^{27,28} (Hi-C) (Box 1)²⁹. Progressively, these techniques were tweaked to allow for
63 enrichment of specific contacts driven by proteins of interest (many-vs-all), including chromatin immunoprecipitation
64 (ChIP-loop³⁰), chromatin interaction analysis with paired-end tag (ChIA-PET³¹), HiChIP³² and proximity ligation
65 assisted ChIP-seq (PLAC-seq³³); or of contacts focused on selected genomic locations (Capture-C²¹⁻²⁵, Capture-Hi-C
66 ²⁶)²⁹

67 Simultaneously with the development of the C-based techniques, ligation-independent techniques were invented
68 to assay not only chromosome conformation in general, but also the nuclear position of chromatin contacts
69 (tyramide signal amplification (TSA), DNA adenine methyltransferase identification (DamID), split-pool recognition
70 of interactions by tag extension (SPRITE)) and **multi-way contacts** (SPRITE and genome architecture mapping (GAM)),
71 which are not assayed effectively using ligation-based techniques^{34–42}. Finally, the recent advancement of super-
72 resolution microscopy and imaging techniques allowed us to investigate chromatin conformation of single cells at
73 extremely high resolution and at a higher throughput than ever before^{12,14,15,43–48}. In addition to improvements in
74 spatial resolution, live-imaging in combination with genome-engineering using CRISPR–Cas9 systems facilitated and
75 improved the study of chromatin-contact dynamics^{49–51}.

76 Owing to these methodological and technological advancements, it is not surprising that the past decade has
77 provided major revelations in 3D genome organization and function. Most notably is the finding that chromosomes
78 in interphase predominantly fold into two compartments, A and B, which respectively consist of predominantly
79 gene-active and gene-inactive regions²⁷ (Figure 1). Furthermore, parts of compartments, from the same or different
80 chromosomes, can come together and create hubs, which are connected by multiple chromatin interactions, thereby
81 sharing a common function (for example, gene repression) and coalescing around different nuclear bodies such as
82 **nuclear speckles**^{35,37,52}. On a scale below the compartments, chromatin interactions were found to be enriched
83 within domains of 100 kb to 1 Mb in length termed topologically associating domains (TADs); these partially insulated
84 domains are subdivided into smaller chromatin nanodomains (CNDs)^{43,53–57} (Figure 1). Both of these layers of
85 organization — compartments and TADs/CNDs — were confirmed to be genomic features present across cell lines
86 and species, but the principles that govern their folding are just beginning to be elucidated^{58–66}. Chromatin-loop
87 extrusion has been shown to be one mechanism responsible for folding at Mb scale. In interphase it is mediated by
88 cohesin complex which can be blocked by CCCTC-binding factor (CTCF) bound to sequence motifs in convergent
89 orientation, thereby demarcating TAD boundaries^{58,59,72–75,61,62,64,67–71} (Figure 1). Importantly, these features are not
90 only of structural nature, but are functional as well, as compartments are quite homogeneously comprised of gene-
91 active or gene-inactive regions and TADs can facilitate the formation of enhancer–promoter contacts within their
92 borders (Box 2; Figure 1). Although it is not entirely clear what the relationship is between TAD boundaries, insulation
93 and disease, structural variations perturbing TAD boundaries and changes in CTCF binding and insulation can alter
94 gene expression and lead to developmental defects and disease^{76–84} (Box 2). These exemplary findings clearly
95 demonstrate the importance of method development and choice in studying 3D genome organization. Below, we
96 discuss established and more recently-developed methods in detail.

97 **2 Sequencing-based techniques**

98 The most common sequencing-based approach to assay chromatin architecture involves crosslinking of spatially
99 proximal chromatin fragments followed by their isolation and sequencing, which is then used as a proxy to estimate

100 contact frequency. Some sequencing-based techniques allow the investigation of chromatin conformation genome-
101 wide (non-enrichment methods), whereas others depend on isolation and identification of a subset of contacts
102 (enrichment methods). Another important distinction is between methods based on ligation of formaldehyde cross-
103 linked chromatin fragments (C-based) and methods that do not involve ligation of cross-linked fragments (non-C-
104 based).

105 **2.1 Non-enrichment methods**

106 Non-enrichment methods assay chromatin conformation at once across entire genomes. The first techniques used
107 ligation to retain together spatially proximal fragments prior to their isolation and sequencing. Subsequently,
108 ligation-free techniques have been developed as well. Both types of approaches capture the majority of 3D features,
109 but they do not deliver identical information due to their inherent technical differences.

110 **C-based methods: Hi-C and Micro-C**

111 Hi-C is frequently used to identify 3D chromatin contacts genome-wide^{27,85} (Figure 2). The original Hi-C protocol
112 included dilution during proximity-ligation, a step introduced originally in 3C and 4C in order to favor ligation of
113 intramolecular chromatin contacts to reduce artifacts. However, this dilution was not very effective, as it was shown
114 that around 60% of contacts originate from inter-chromosomal interactions⁸⁵⁻⁸⁷. This problem was circumvented by
115 the omission of the SDS treatment just before ligation, which allowed permeabilization of the nuclear membrane
116 and thus chromatin digestion and ligation *in situ*^{85,86}. This modification is justified by the fact that ligation occurs in
117 fixed nuclei, where molecular diffusion is virtually absent, and by the fact that the cross-linking reaction occurs
118 between reactive moieties that are located in the nanometer range. Therefore, *in situ* ligation enabled a more
119 efficient capture of true contacts, thereby delivering higher resolution for the same sequencing depth. After this
120 methodological improvement, the omission of SDS treatment was adopted in virtually all subsequent C-based
121 techniques. *In situ* chromatin processing allowed further development of single-cell Hi-C, which was the first
122 sequencing-based single-cell chromatin analysis technique to be successfully established⁸⁷.

123 Although Hi-C is suitable for the detection of compartments and TADs, its resolution is intrinsically linked with the
124 use of restriction enzymes and to sequencing depth. Therefore, even if Hi-C libraries are sequenced by the billions
125 of reads, the unbiased detection of local interactions, like enhancer–promoter contacts in the sub-TAD range, largely
126 depends on the distribution of restriction sites, which in turn depends on the underlying sequence composition and
127 thus is not uniform across the genome. This issue was first addressed concomitantly in two techniques, DNase Hi-C
128 and Micro-C, both based on chromatin fragmentation without using restriction enzymes⁸⁸⁻⁹¹. Micro-C introduced
129 double cross-linking and replaced the restriction enzymes used in Hi-C with micrococcal nuclease digestion⁹¹⁻⁹⁴
130 (Figure 2). This produces a fairly uniform fragmentation down to the nucleosome level, which increases local
131 resolution. In addition, Micro-C (and theoretically DNase Hi-C) also retains information on nucleosome positioning,
132 which can be jointly analyzed with chromatin contact information from a single data set. However, although this
133 technique is a promising improvement for the study of local chromatin topology, according to a recent preprint

134 article, it is less efficient in capturing long-distance and inter-chromosomal contacts compared with Hi-C⁹⁵.
135 Therefore, careful framing of the research questions is needed to accurately select between different non-
136 enrichment methods.

137 **Non-C-based methods: SPRITE and GAM**

138 Although C-based methods have been extremely successful; they have intrinsic limitations and potential sources of
139 bias. First, like many other chromatin analysis methods, they rely on mild formaldehyde cross-linking, which is
140 powerful but potentially limited in capturing interactions of proteins with short residence time on chromatin or
141 containing a low fraction of amino acids that can be crosslinked^{96,97}. Second, these methods require ligation of
142 genomic fragments prior to sequencing, a procedure that is only partially efficient. Third, they depend on short
143 paired-end sequencing, which provides information only on bipartite interactions, whereas multipartite *in vivo*
144 chromatin interactions escape the analysis.

145 Several methods that do not rely on ligation of chromatin fragments allow the detection of dual or multiple
146 interactions. In SPRITE, crosslinked nuclei are isolated and fragmented, then individual crosslinked pieces of
147 chromatin are uniquely barcoded using multiple cycles of a split-and-pool strategy; after high-throughput
148 sequencing, reads carrying the same combination of barcodes represent genomic sites that are a part of the same
149 crosslinked cluster (Figure 3). Since the method does not select for specific sequence sizes, it yields bipartite as well
150 as multipartite contacts. SPRITE has been further adapted to facilitate the capture of DNA–DNA, RNA–DNA and RNA–
151 RNA interactions, thereby allowing to determine whether RNAs of interest are associated with a subset of genome
152 interactions and what the relationship of the RNA is to nuclear landmarks³⁷. Finally, a single cell version of SPRITE
153 allows the study of multi-way contacts in individual cells³⁶. Future work applying this method to a broad range of
154 cell types and analyzing two-way and multi-way contacts to a deeper level is required in order to fully exploit its
155 advantages and appreciate its limitations⁵². GAM is an orthogonal method that can also provide frequencies of
156 multivalent interactions³⁵. In GAM, fixed cells are embedded in sucrose, frozen and cryo-sectioned, and the DNA is
157 extracted and sequenced from each section^{41,98} (Figure 3). Loci that are closer to each other in the nuclear space are
158 co-sequenced more frequently than distant loci. Since sections are taken from multiple nuclei sliced at random
159 orientations, the co-segregation of all possible pairs of loci among a large collection of nuclear section profiles is
160 used to generate a matrix of inferred locus proximities. GAM matrices produce maps similar to Hi-C maps but require
161 fewer cells — few hundred nuclei produce maps that approximate those obtained from large populations of cells in
162 Hi-C. Like SPRITE, GAM can identify multiple interactions, thereby enabling the direct study of multivalent enhancer-
163 promoter interactions and of higher-order chromatin structures.

164 **2.2 Enrichment methods**

165 The above-described techniques detect chromatin contacts present in the nucleus irrespective of genomic location,
166 nuclear topography or the underlying protein binding. Yet, to fine-scale map chromatin folding and understand some
167 of its functional aspects, it is necessary to detect specific contacts using enrichment approaches, thereby amplifying

168 the contacts signal in a specific genomic region of interest (Capture-C, capture Hi-C (cHi-C)) or for a specific protein
169 of interest (ChIA-PET, ChIP-loop, HiChIP, PLAC-seq, DamID, DamC, tyramide signal amplification (TSA-
170 seq))^{21,22,38,39,42,99–101,23–26,30,32–34} (Figure 2).

171 **C-based methods: HiChIP, Capture-C and cHi-C**

172 The first C-based enrichment-dependent techniques to be developed combined proximity-ligation with ChIP, for
173 example the low-throughput ChIP-loop and later, with higher-throughput, ChIA-PET, which was subsequently
174 improved to allow for more efficient mapping and detection of single-nucleotide polymorphism^{30,31,102}. However,
175 proximity-ligation in these techniques was performed in non-optimal conditions, with ChIP and sonication done
176 preceding ligation, possibly affecting the accuracy of captured interactomes. These shortcomings were addressed in
177 the next generation of protein-enrichment C-based techniques, HiChIP and PLAC-seq, in which the C-technique is
178 performed first and *in situ*, thereby creating an optimal environment for proximity-ligation, followed by protein
179 enrichment^{32,33,102}. HiChIP and PLAC-seq allow the identification of a subset of interactions forming in association
180 with specific protein binding, but can only be used as a proxy, not a definitive proof, that a protein of interest
181 mediates the captured chromatin contacts. In addition to protein-mediated enrichment, chromatin contacts can also
182 be enriched for a specific genomic location(s) using techniques such as Capture-C and cHi-C^{21–26} (Figure 2). In these
183 techniques, a Hi-C library is first generated and then hybridized to specifically designed baits (RNA or DNA)
184 corresponding to either one large region (several Mbs) of interest in the case of cHi-C, or to multiple specific sites in
185 the genome (for example, a collection of specific gene promoters) in the case of Capture-C. This allows mapping of
186 contacts in fine detail, which would normally require 20–50 fold more sequencing without enrichment^{26,76,103}. In the
187 future, capture approaches could be combined with other techniques to adjust them for specific needs.

188 **Non-C-based methods: TSA and DamID**

189 Ligation-independent techniques such as TSA-Seq and DamID can also enrich for contacts associated with specific
190 proteins and map the nuclear topology^{34,38,39,42,100}. TSA-Seq relies on **tyramide signal amplification**^{34,42}. Cells are first
191 crosslinked, stained with a primary antibody against a protein of interest and then with a horseradish peroxidase
192 (HRP)-conjugated secondary antibody^{34,42}. HRP catalyzes the formation of tyramide–biotin free radicals, which
193 diffuse and covalently link to nearby proteins, DNA and RNA. The biotin moiety can be used to stain nuclei as well as
194 purify and sequence the associated DNA. Since the amount of tyramide–biotin signal decreases with increasing
195 distance from the antibody localization source, TSA-Seq read frequencies can be transformed into 3D distances from
196 nuclear landmarks of interest upon appropriate calibration. A second and orthogonal, cross-linking-independent
197 technique is DamID, which involves tethering *Escherichia coli* DNA adenine methyltransferase (Dam) to a chromatin
198 protein; the Dam moiety methylates adenines at GATC consensus DNA sites surrounding the chromatin protein of
199 interest³⁸. Application of DamID to proteins with distinct nuclear compartmentalization allows identifying genomic
200 domains associated to nuclear landmarks, such as the **lamina associated domains (LADs)**⁴⁰. More recent applications
201 of DamID also enable identifying LADs in single cells and to simultaneously quantify protein–DNA contacts and RNA
202 expression in the same cell^{100,104}. Similar to DamID, DamC has been established as a cross-linking and ligation free

203 technique that can replace 4C; in DamC, a fusion protein of Dam and reverse tetracycline receptor (rTetR) is recruited
204 to Tet operator sequences (TetOs) ectopically inserted at a genomic site of interest¹⁰¹. Methylated DNA is then
205 detected by high-throughput sequencing, and scoring of the methylated Dam target sites around the TetOs allows
206 quantifying chromatin contacts. DamC may be of great interest for low cell number or tissue-specific applications.

207 3 Super-resolution microscopy methods

208 In DNA FISH, DNA probes are hybridized to cognate genomic regions of interest and visualized by fluorescence
209 microscopy, which allows measuring localization, shape and inter-probe distances⁴⁻⁶. However, the study of
210 chromosome conformation and of individual chromatin contacts under the microscope has been limited by the low
211 number of loci that can be probed simultaneously owing to the low number of available independent fluorescence
212 channels, and by the limited spatial resolution of traditional light microscopes. These limitations have been removed
213 thanks to major technological advancements in light microscopy applications.

214 Light emitted by any point source is diffracted such that the point will appear in an image as a so-called **airy**
215 **diffraction pattern**, the size of which is proportional to the wavelength. In practice, this property of light, together
216 with aberrations of optical systems and light scattering, limits the resolution (the minimal distance at which two
217 signals can be distinguished), following a formula derived by the physicist Ernst Abbe, to approximately 250 nm in
218 the x and y axes (lateral resolution) and 600 nm in the z axis (axial resolution), even when using the best confocal
219 microscopes and image processing software. For decades, the resolution of light microscopy was believed to be
220 intrinsically limited by diffraction, but imaging technologies have progressed at a remarkable speed, allowing the
221 detection of increasing number of nuclear components at a spatial resolution surpassing the Abbe limit. These
222 methods, collectively called super-resolution microscopy, increase spatial resolution mainly in three different ways,
223 and here we discuss their applications in 3D genome organization research (Figure 4a). Detailed descriptions of these
224 methods can be found in more specialized microscopy reviews^{105,106}.

225 Structured illumination microscopy (SIM) is an approach to super-resolution microscopy that increases resolution by
226 a factor of two in each axis by exploiting a non-uniform illumination pattern: the sample is serially exposed to light
227 from different angles and different axial phases¹⁰⁷⁻¹⁰⁹ (Figure 4a). This illumination pattern interferes with the
228 sample in a manner that can be conveniently analyzed in the Fourier mathematical space to improve resolution.
229 Despite its complexity, the advent of commercially available SIM microscopes and software and the fact that the
230 technology is compatible with standard fluorophores, labeling procedures and multi-color imaging has stimulated
231 the widespread application of SIM. SIM has allowed to assess unprecedented details of chromatin and nuclear
232 organization, such as the relations between chromatin and the nuclear periphery or the physical structure of TADs
233 in single cells^{43,53,110}. In particular, the analysis of mammalian TADs using SIM has revealed that TADs are subdivided
234 into the smaller CNDs: as CND boundaries vary from cell to cell, ensemble Hi-C experiments blur their identification,
235 thereby illustrating the power of single-cell, super-resolution imaging to illuminate 3D genome organization features
236 that are inaccessible to cell-population-based technologies^{57,111}.

237 A second family of super-resolution microscopy methods is called single molecule localization microscopy (SMLM);
238 it includes stochastic optical reconstruction microscopy (STORM), photo-activated localization microscopy (PALM)
239 and fluorescence photo-activated localization microscopy (FPALM) ^{112–114}. All three methods use fluorophores that
240 can be converted from a fluorescent (or activated) state to a dark (or inactivated) state (and vice versa) and rely on
241 the stochastic excitation and detection of spatially separated single fluorophores. The spatial separation is achieved
242 by making sure that only a small fraction of the total population of fluorophores in the sample can emit light, so that
243 individual emitters do not overlap in a single imaging frame ¹¹⁵. Sequential imaging using cycles of activation and
244 inactivation of the fluorophores, followed by the precise localization of the center of emission of the individual
245 fluorophores and on the superposition of all imaging cycles, generates super-resolution images that can reach a
246 lateral resolution of 20 nm in biological samples ((3D images can be obtained using various methods)¹¹⁵ (Figure 4a).
247 The application of these methods has provided crucial insights into the fundamental folding of chromatin in the
248 nucleus. Nucleosomes were shown to transiently interact to form clutches of various sizes interspersed with
249 nucleosome-depleted regions ¹¹⁶. At a higher scale of organization, nucleosomes were found to form CNDs, i.e.,
250 aggregates of a diameter of ~160 nm, within which individual nucleosomes display highly correlated motion in live
251 cells ¹¹⁷. This organization suggests that CNDs may arise from a coordinated behaviour that might reflect multiple,
252 dynamic nucleosome interactions, consistent with a recent analysis of fixed chromatin by SIM ^{57,111}. It will be
253 interesting to study whether CNDs might regulate genome functions such as gene expression (in particular, the
254 frequency of Enhancer-Promoter contacts) or DNA replication.

255 The third approach to super-resolution microscopy is stimulated emission depletion (STED). This technique uses a
256 configuration similar to confocal microscopy, but with an additional laser called the depletion beam, which
257 illuminates the sample in a donut shape that has zero intensity at the center of the excitation laser ^{118,119}. This
258 illumination provokes the depletion of emissions in the periphery and only allows emission in the center, thereby
259 generating a **sub-diffractive point spread function**. In most applications, STED reaches a lateral resolution of 30–50
260 nm and high axial resolution can also be obtained (Figure 4a). The drawbacks of this method are its strong laser
261 intensity and the requirement for specific fluorophores. A modified version called reversible saturable optical
262 fluorescence transitions (RESOLFT) allows using lower beam intensities ¹²⁰. Importantly, SMLM and STED do not have
263 a theoretical resolution limit and a combination of the two methods has achieved axial and lateral resolutions below
264 3 nm in cells ¹²¹. Therefore, imaging applications are quickly enabling to investigate chromatin and nuclear
265 architecture at the macromolecular scale.

266 A remaining limitation of these super-resolution microscopy methods is throughput, as they are limited to using
267 two or three colors, which restricts the number of loci that can be analyzed simultaneously. Recently, however,
268 oligonucleotide-based FISH protocols called oligopaints were coupled with microfluidics to allow multiple cycles of
269 hybridization¹²² (Figure 4b, 4c). This conjunction of methods enabled sequential probe hybridization and assaying
270 multiple different loci with high precision in the 3D nuclear space ^{12,15,44–48,123}. These techniques allowed obtaining
271 distance distribution maps among all imaged points that recapitulate the maps obtained from Hi-C experiments and

272 provide 3D trajectories of chromatin molecules at high resolution and in thousands of cells, something which is
273 unattainable even in single-cell Hi-C (Figure 4c, 4d). These techniques include Multiplex FISH imaging, which helps
274 establish high resolution tracing of chromatin folding of megabase-size genomic domains by labeling several tens of
275 locations in the genome simultaneously^{44,46,48,123} optical reconstruction of chromatin architecture (ORCA)⁴⁷; Hi-M
276⁴⁵, which is a multiplexed, sequential imaging approach; and oligopaint fluorescence *in situ* sequencing (OligoFISSEQ)
277¹⁴ (Figure 4b, 4c, 4d). Importantly, OligoFISSEQ is a method combining barcoded Oligopaint to *in situ* sequencing
278 technology, which is used to read out the barcode carried by the oligos (Figure 4d). This makes OligoFISSEQ a
279 powerful high-multiplexing technology¹⁴. *In situ* sequencing can also be coupled to Tn5 transposase-mediated
280 random insertion of DNA-sequencing adapters into hundreds of positions of fixed genomic DNA. *In situ* amplification
281 can then be used to insert unique molecular identifiers that are sequenced *in situ* prior to *ex situ* sequencing in order
282 to identify the genomic region of Tn5 insertion. This allows obtaining the 3D location of hundreds of loci per cell, a
283 powerful alternative to hybridization-based imaging methods¹⁶.

284 Currently, it is possible to simultaneously visualize thousands of DNA loci, hundreds of different RNA molecules
285 and several proteins or histone modifications, thereby enabling high-throughput structure–function analyses in
286 thousands of single cells and truly inaugurating the field of spatial imaging-based 3D genomics^{12,15}. These light
287 microscopy methods are accompanied by developments in electron microscopy, with electron-microscopy
288 tomography allowing the study of chromatin at nanometer resolution¹²⁴. Each of these methods has advantages
289 and limitations. For instance, methods with very high spatial resolution are typically not optimal for the description
290 of architectures of large domains owing to a slow acquisition process and to the intrinsic noise in the images that
291 are obtained. However, advanced OligoSTORM imaging provides powerful information on 3D genome organization
292 that is complementary to molecular techniques such as Hi-C, thereby enabling the investigation of genome
293 architecture and function to a degree that was unthinkable a decade ago¹⁴.

294 **4 Computational analysis and modelling**

295 Although the methods discussed above provided important insights into 3D genome organization and function, they
296 are still limited in their ability to describe how the chromatin fiber folds in the 3D space of the nucleus and they
297 cannot predict structural changes that would result from perturbations such as mutations in genes or in gene-
298 regulatory components. Evaluating the impact of architecture on genome function remains even more inaccessible
299 at present. All these limitations have stimulated computational analyses and the development of mathematical
300 modelling which, in conjunction with experiments, might help achieve a quantitative and predictive understanding
301 of chromosome architecture and function.

302 4.1 Analysis of Hi-C data

303 The advent of Hi-C and related technologies has raised strong interest in the development of matching
304 computational analysis tools, owing to the inherent complexity of Hi-C data. The achievable spatial resolution of Hi-
305 C is affected by sequencing depth, library complexity and the DNA-cutting frequency of the enzyme used for
306 chromatin fragmentation. Since the number of possible chromatin-fragment interactions is extremely high ($> 10^{14}$ in
307 the human genome, when using 4-base cutter restriction enzymes) and the sequencing depth of a typical experiment
308 is limited, Hi-C matrices are sparse, that is, many entries in the matrices are 0, if they are not represented at the
309 appropriate resolution. This makes it impossible to distinguish between genuine absence of contacts and absence
310 of contacts owing to low sequencing depth. Furthermore, the different sizes of restriction fragments across the
311 genome, differences in mappability between regions with high or low density of repetitive elements, and the decay
312 of interaction frequencies with increase in genomic distance make Hi-C matrices typically very heterogeneous in
313 terms of contacts at different genomic locations or across different distances. All these factors limit the resolution
314 with which one can call contact regions or domain boundaries, and generate difficulties in defining the precise
315 locations of compartments, TADs and chromatin loops. The first Hi-C study²⁷, which used a 6-base cutter and
316 achieved low sequencing depth, produced reliable matrices at the resolution of 100kb and identified the
317 compartments. To this end, Hi-C matrices were normalized by genomic distance, converted to correlation matrices
318 and subjected to principal component analysis, which distinguished the active (A_ and inactive (B) compartment
319 types. Later algorithms additionally applied clustering steps like Gaussian hidden Markov modelling to Hi-C maps
320 based on much deeper sequencing, leading to further specification of epigenetic compartment signatures and more
321 detailed stratification of the A compartment into two sub-compartments and B compartment into three sub-
322 compartments^{85,125–129}.

323 The computational identification of TADs required high resolution maps, which were published three years after the
324 first low-resolution HiC experiment^{53–55,130}. Although nowadays TAD calling is done routinely, there is no clear
325 consensus method but rather numerous TAD callers that are based on different principles. Initial computational
326 approaches such as insulation score and directionality index determined TAD boundaries by defining a one
327 dimensional linear score of a bin-fractionated genome, where the bin of the local minima (for the insulation score)
328 or the one between local minima and maxima (for the directionality index) would determine the boundary position
329^{55,131,132}. These approaches, however, could not inform on TAD hierarchy and missed identifying nested TADs.
330 Subsequently, other computational approaches were developed to address this issue either by further developing
331 the linear score approach (Matryoshka¹³³), by clustering contact map data (ICFinder¹³⁴, TADpole¹³⁵) or by using
332 graph theory-based algorithms that identify nested TADs as contact subnetworks connecting to form larger TADs
333 (3DNetMod¹³⁶). Furthermore, numerous other computational tools have been developed that combine these
334 approaches with different efficiencies at different scales and resolutions^{137–140}.

335 The third major feature of Hi-C data are chromatin contacts and loops, which, like TADs, became detectable as the
336 data became more resolute. Specific chromatin contacts are defined as statistically significant increases in contact

337 detection in comparison with a general background model. This is the basis of Fit-Hi-C, a computational tool that
338 assigns a statistical confidence to a contact by using random polymer modeling while accounting for known Hi-C
339 biases such as genomic distance; an adapted version, HiC-DC, additionally accounts for sparsity and over-dispersion
340 and yields a more conservative statistical significance estimate^{141,142}. However, a locally enriched contact — a
341 chromatin loop — might elude algorithms that use only a general background model to estimate statistical
342 significance of enrichment. HiCCUPS, one of the first loop-dedicated algorithms, identifies a chromatin loop as the
343 most enriched bin in comparison to its immediate neighborhood while using a high-resolution 5kb Hi-C data as an
344 input⁸⁵. This algorithm helped drive the discovery of a specific subtype of loops, CTCF loops, and contributed to the
345 development of the loop-extrusion model, thereby demonstrating the importance of specialized algorithm
346 development in order to grasp the full biological significance of the experimental data. In-depth reviews and
347 websites have compiled and compared available compartment, TAD and loop callers^{139,143–147}. The increasing
348 robustness of primary computational tools to identify 3D genome features has led to a blooming of applications
349 aimed to identify DNA sequences of biological significance linked with genomic regions in spatial proximity as
350 mapped by Hi-C techniques. These applications have enabled associating genes with putative enhancers based on
351 contact frequency, epigenomic and DNA features; identifying novel regulatory elements from genome-wide
352 association studies; and assigning a potential role for short tandem repeats in genome organization^{148–155}.
353 Although these computational tools can provide precious information on genome structure and function, the cost
354 of obtaining high-resolution Hi-C maps can become prohibitive, particularly when many experimental conditions
355 need to be compared. In order to tackle this problem, a machine learning approach based on deep convolutional
356 neural networks has been used to impute higher-coverage Hi-C maps from low-coverage data in order to increase
357 the resolution with which loops or TAD borders can be defined¹⁵⁶. This function is important in order to identify
358 genomic features that might be involved in the regulation of these structures even if the data is not sufficiently
359 resolute.

360 **4.2 3D chromatin modelling**

361 The computational tools discussed above provide new information about regulatory elements and their function,
362 but they do not inform on 3D architecture. Three main types of modelling strategies have been used to infer 3D
363 genome folding, although some methods can blur this classification as they have characteristics belonging to more
364 than one type^{137,157–160}. The first modelling strategy is the bottom-up approach of polymer modelling, which
365 attempts to infer and understand chromosome architecture from first principles, typically modelling chromosomes
366 as self-avoiding polymers moving in a confined space that represents the nucleus (Figure 5a). Applications aim to
367 identify components regulating 3D folding, and thus they try to reduce the number of variable parameters describing
368 the polymer behavior while maximizing the fit between experimental data, typically Hi-C contact matrices, and
369 analogous matrices that are derived from measuring contacts in snapshots taken at given times after starting
370 polymer motion simulations. This modelling strategy has been applied to rationalize the decay in contact frequencies

371 between chromatin regions as a function of the linear distance separating them^{161,162}. More recently, these models
372 have offered an explanation for the formation of domains of active and inactive chromatin, the generation of TADs
373 through loop extrusion and the contribution of epigenetic features such as chromatin types to the formation
374 of TADs and compartments^{58,67,128,163–166}. The combined role of loop extrusion and of active and inactive
375 compartments in the determination of global chromosome organization has also been studied⁶⁸. Furthermore,
376 in addition to intra-chromosomal contact frequencies, this strategy allowed the investigation of inter-
377 chromosomal contacts¹⁶⁷. A current limitation of these models is that they can typically reproduce and predict
378 some, but not all the features of 3D chromosome folding and in particular, they usually do not perform equally
379 well at different scales (loops, TADs, compartments, chromosome territories)^{137,138}. This is partly due to the
380 considerable computational time required for the iterative simulation processes that are involved in generating
381 the models, and interesting ongoing developments involve accelerating computation¹⁶⁸.

382 An alternative, physics-based bottom-up modelling strategy does not aim at minimizing the number
383 of parameters to describe polymer behavior. Instead, it models chromosomal regions as polymers, in which
384 each monomer represents a genomic region of fixed size and can interact with any other monomer with a
385 specific energy. Each of the interaction energies can be adjusted until the configuration ensemble of the
386 polymer produces a contact matrix that resembles the Hi-C data matrix. This approach allows searching for the
387 monomers that have the most crucial role in driving the specific 3D configuration defining the genomic region
388 of interest^{169–171}. Furthermore, it can also be used to study the 3D path of the chromatin fiber in the resulting
389 polymer models and to compare it with 3D data such as those provided by imaging methods in order to relate
390 contacts to 3D architectural features of the region of interest⁴³.

391
392 The second type of modelling strategy is the top-down approach of restraint modelling, starting from data,
393 which are often derived from Hi-C maps and sometimes integrated by maps of chromatin
394 –nuclear lamina interactions, in order to infer the 3D architecture of genomic loci, entire chromosomes or the
395 whole nucleus (Figure 5b). In some applications, the contact maps are used to set restraints that the models
396 must satisfy in setting the 3D folding paths of chromatin fibers. The modelling result can be a consensus
397 genome structure, or an ensemble of structures, which reflects the structural variability among cells or during
398 time^{28,172–179}. This type of modelling can also deliver information about chromatin folding dynamics, provided
399 Hi-C data from time-course experiments are available. This is possible by interpolating the restraints through
400 each of the time points¹⁸⁰. Another interesting data-driven modelling uses a population deconvolution
401 approach, in which Hi-C data are used to generate a large population of structures which, together, reproduce
402 the experimental interaction patterns. This approach has been extended to incorporate chromatin–lamina
403 interaction and imaging data^{46,181,182}.

404 The recent progress in computation speed led to the deployment of another cohort of 3D chromosome folding
405 prediction tools, which are based on machine learning methods that use epigenomic and chromosome

406 conformation information as input and display the predicted 3D architecture as output ^{171,183–188}. Such
407 computational tools rely on input data obtained from several different cell lines to train their models and identify
408 the minimum necessary signature to accurately predict an enhancer–promoter pair, promoter–promoter pair and
409 CTCF loops as well as for contact quantification ^{183,184,187,188}. Recently, two tools, Akita and DeepC, used convolutional
410 neural networks to predict 3D folding solely on the basis of DNA sequence ^{185,186}. These kinds of computational tools
411 could become very important to enable making predictions from experimental samples, for which the full epigenome
412 datasets are not available because of sample-quantity limitations, for example in the case of patient samples.

413

414 A third type of modelling strategy, which combines both top-down data driven models and bottom-up physical
415 models has been described ¹⁸⁹. This model uses parameters derived entirely from a Hi-C experiment as input,
416 but also factors in a polymer’s energy function like in bottom-up approaches. However, in this case the energy
417 function is designed strictly from biological factors that have been demonstrated to have a role in 3D genome
418 organization, which ensures biological relevance while simultaneously allowing for mechanical investigation.
419 With monomers of 1 Mb in size, this model successfully reconstituted the radial positioning of entire
420 chromosomes and uncovered previously unknown contributions of distinct biological processes (separation of
421 A and B compartment, centromere clustering, inter-chromosomal contacts). However, the large monomer size
422 limited resolution and prevented the capture of more local features such as sub-TAD or loop structures, but
423 the currently available computational power could allow decreasing the monomer size and testing whether
424 mechanistic insights into more local structures can be correctly modelled.

425 Improved algorithms and the ever-increasing computation power will soon allow modelling the dynamics of
426 whole-genome folding at high spatial and temporal resolution, making computational methods crucial
427 complementary tools to the experimental methods.

428 5 Emerging genome structure technologies

429 Many outstanding questions remain in the research of nuclear architecture, and it is therefore not surprising that
430 new sequencing-based methods, microscopy-based methods and computational methods are continually being
431 developed.

432 One outstanding question in the field is how to address single-cell variability while not compromising high-
433 throughput. Sequencing-based techniques address this through the adaptation of C-based techniques and C-
434 independent techniques for single cells ^{36,87,190}. The first of these single-cell adaptations was single-cell Hi-C (scHi-C),
435 which revealed high inter-cell contact and TAD variability and indicated that TADs are highly stochastic domains ⁸⁷.
436 This finding, however, put into question whether this major Hi-C feature, TADs, represents actual physical structures
437 or reflects statistical average rather than a physical reality. Later, microscopy studies settled this controversy by
438 showing that TADs do correspond with physical domains, but also that they have highly variable structures, clearly

439 emphasizing the need to focus on techniques that provide information on large numbers of single cells^{43,44,57}. Indeed,
440 an increasing number of chromatin analysis techniques are being developed into single-cell applications to address
441 this issue and study chromatin conformation stochasticity and inter-cell variability^{36,190–193}.

442 Another major goal pertaining to C-based methods, is to overcome the resolution limit and potential cross-linking
443 bias. Although alternative fragmentation techniques (Micro-C, DNase-Hi-C) successfully deal with the resolution
444 problem, every technique that relies on crosslinking is inherently biasing fragmentation towards open-chromatin
445 regions^{88,91}. A recently published method, Cap-C, approaches this problem by exchanging the standard
446 formaldehyde crosslinking with **dendrimer crosslinking**. By using three differently sized molecules, Cap-C allows
447 homogeneous crosslinking of open and closed chromatin, thereby achieving more uniform fragmentation and higher
448 resolution^{88–91,194}. Since this is a straightforward change to current C-based protocols, it has the potential to be
449 widely implemented.

450 The question of how to obtain spatial coordinates of the chromatin in the nucleus in a high-throughput manner and
451 how to integrate chromosome conformation data with the spatial position has been difficult to answer using
452 sequencing-based techniques. Recently, two new methods to study nuclear topology and higher-order organization
453 based on ligation-free methodology were published, genomic loci positioning by sequencing (GPSeq)¹⁹⁵, and
454 chromatin interaction analysis by droplet based genomic analysis (ChiA-Drop)¹⁹², which provides information on
455 multivalent interactions (similar to SPRITE). In ChiA-Drop, cross-linked and fragmented chromatin is loaded onto a
456 microfluidics device so that individual cross-linked molecules are partitioned into droplets that contain unique
457 barcoding reagents (Figure 3). After pooling, high-throughput sequencing and identification of reads carrying the
458 same barcodes, putative 3D interactions are identified. In *Drosophila melanogaster*, ChiA-Drop was performed using
459 less than 10,000 cells and thus it could be suitable for analyzing rare cell types. Furthermore, the possibility to enrich
460 for interactions that depend on specific proteins allows to infer the relative position of the interacting regions
461 relative to nuclear bodies or landmarks¹⁹². GPSeq primarily focuses on the study of radial chromosomal positions in
462 the nucleus by performing restriction enzyme digestion in a time course *in situ*, which allows the capture of the most
463 nuclear-periphery-adjacent chromatin following short-term digestion, whereas the longer digestion times capture
464 progressively more interior parts of the genome. However, to infer distances to the periphery correctly it is necessary
465 to perform a YFISH (in which a Y-looking FISH adapter is ligated on the restriction enzyme overhang while the other
466 side of the adapter interacts with FISH probes) coupled with super-resolution imaging in a time course. Using this
467 approach, it is possible to investigate not only the radial position of chromosomes but also the radial positions of
468 DNA replication, double-stranded DNA breaks (DSBs) and mutations¹⁹⁵. Recently, a computational method called
469 SPIN (spatial position inference of the nuclear genome) has been developed to predict genome-wide spatial
470 positioning in the nucleus. The method integrates spatial multi-omics data including TSA-seq, DamID and Hi-C in a
471 computational framework based on Hidden Markov random field to localize clusters of chromatin contacts relative
472 to nuclear bodies such as nuclear speckles or the lamina¹⁹⁶. This complementary tool of experimental and

473 computational multi-omics methods might provide the essential missing components in the nuclear organization
474 research toolbox.

475 Finally, the last outstanding questions in the field we shall discuss are how to achieve higher spatial resolution in
476 microscopy, and how to assay chromatin dynamics of individual genomic loci. In order to achieve spatial resolution,
477 one has to be able to image beyond the **diffraction limit**. In addition to super resolution microscopy, several recent
478 publications reported an alternative approach called expansion microscopy (ExM), in which the sample is embedded
479 in a polyelectrolyte gel that expands 4–5 times when immersed in water^{197–199}. ExM offers imaging of structures
480 that are beyond the diffraction limit using conventional microscopy, and according to a recent preprint article, when
481 combined with super-resolution microscopy, ExM achieved resolution of 5 nm¹⁹⁸. Furthermore, ExM can be
482 extremely powerful for spatially precise positioning of RNA species *in situ* and has been recently used in combination
483 with FISSEQ technology to perform RNA *in situ* sequencing in an unbiased manner²⁰⁰. This technology
484 conglomeration offers great promise for the future, as it combines different principles to achieve sub-diffractive
485 resolution and multiplexing. One could easily imagine extending its applications to multiplex DNA FISH or *in situ*
486 sequencing methods in order to analyze the traces of chromosomes, chromosome domains or individual loci at high
487 resolution. However, it is important to keep in mind that sample expansion can alter the ultrastructure of the
488 chromatin and it will be necessary to ensure that the structure remains preserved under standard conditions.

489 There are two major bottlenecks to studying chromatin contact dynamics of individual loci. First, in order to visualize
490 such contacts, cells usually had to be subjected to heavy genome engineering to insert either lac-O or tet-O arrays
491^{50,201–203}. Second, the signal must be sufficiently strong in order to visualize individual loci. Multiple different methods
492 have been developed to reach this goal^{49,50,204–207}. Chimeric array of gRNA-oligo (CARGO) and CRISPR–Cas-mediated
493 Live FISH are examples of two independently developed live-imaging techniques that addressed these limitations by
494 using the CRISPR–Cas9 system. CARGO uses multiple guide RNAs (gRNAs) targeting nuclease-dead Cas9 (dCas9)-
495 EGFP to certain genomic locus to achieve efficient fluorescence-signal amplification and circumvent the need for
496 strenuous genome engineering⁴⁹ (Figure 6a). This method enabled the detection of *cis*-regulatory element mobility
497 during cell differentiation in relation to their expression, and offers great promise as it is relatively simple⁴⁹. Similarly,
498 CRISPR–Cas-mediated Live FISH utilizes dCas9 to target a region of interest, but here the gRNAs are fluorescently
499 labeled, thereby amplifying the signal more than fourfold²⁰⁶ (Figure 6b). Furthermore, the use of catalytically active
500 Cas9 together with dCas9 allowed the simultaneous visualization of DSBs and fluorescent DSB-repair proteins,
501 practically creating a live Immuno-FISH²⁰⁶. Finally, Live FISH was further expanded by coupling the dCas9–gRNAs
502 with the dCas13–gRNA system, thereby granting visualization of both DNA and nascent RNA transcripts in live cells
503^{206,207}. Together, these and similar dCas9-based techniques might become valuable for studying chromatin and
504 transcription dynamics in live cells, and open venues for application ranging from basic science to diagnostics^{50,208}.

505 **6 Future directions**

506 Genome architecture as a field of research has come a long way in a remarkably short period of time thanks to multi-
507 disciplinarity that was driven by technological advancements. In this Review, we discussed major discoveries in
508 chromatin conformation and nuclear topology through a technical prism. However, there is still so much we do not
509 understand and that is not accessible to us owing to methodological shortcomings.

510 The next decade will likely see a continued expansion of imaging-driven techniques with a strong emphasis on
511 multiplexing and on live microscopy, especially in conjunction with sub-diffractive resolution. As already now we see
512 the implementation of Live FISH, further development and specialization of these methods could possibly help to
513 study enhancer–promoter dynamics in respect to transcriptional output^{206,207}. Furthermore, live-microscopy-based
514 techniques could be well suited to study the kinetics of transcription factor binding to chromatin, a subject that is
515 poorly understood. In addition to live microscopy, the throughput of the super-resolution FISH-based methods will
516 likely further increase and help study inter-cell variability. However, these predictions are based on the premise that
517 the hardware necessary for these techniques will become more available and affordable.

518 Regarding sequencing-based methods, there is already a strong tendency to use different methods on the same
519 sample, for example nucleosome occupancy and Hi-C (Micro-C, DNase-Hi-C) or bisulfite sequencing and Hi-C, in order
520 to obtain different types of information from the same sample, but also to expand the amount of information that
521 can be extracted from limited material such as patient or rare-cell samples^{91–93,191}. This trend will likely continue,
522 and new multifaceted approaches will emerge allowing the collection of complex data. Furthermore, existing
523 techniques are constantly being adapted to extend their applicability. For instance, two recent Hi-C-based
524 techniques and a SIM imaging-based technique have been described that allow inspection of sister chromatid
525 topology at the cell population level^{73,74,209}. These techniques mark a significant milestone because, unlike other
526 sequencing-based techniques, they allow the study of chromatin conformation during S phase of the cell cycle. It is
527 possible that the implementation of these techniques will fuel discoveries related to S phase and DNA replication
528 that were unattainable with the previously available techniques.

529 An important open question is the causality between topological insulation and transcription, which currently is
530 difficult to properly address. Ideally, a single-locus proteomics approach would be appropriate to investigate the
531 underlying proteome of a TAD border or a local chromatin insulator region in order to identify candidate insulator
532 factors and analyze whether they play a causal role to determine insulation. Existing single-locus proteomics
533 techniques to study topological insulation and transcription are available, but they are incredibly laborious and
534 complicated to implement^{210–212}. A major breakthrough in proteomics or in wet lab protocols to decrease input
535 material is required for single-locus proteomics techniques to become widely applicable. However, it is an exciting
536 lane of research that will certainly help explain the functional aspect of chromatin conformation.

537 Finally, we argue that not only methods driven by a technological boom, but also different and unconventional points
538 of view should coalesce to invent new approaches and fuel milestone discoveries in genome architecture research.
539 The importance of the interdisciplinary approaches described above will become even more prominent with future
540 technological developments. Indeed, in order to re-orientate our field of research in the interdisciplinary direction,
541 large consortia are being organized with the purpose of connecting different expertise and points of view. We believe
542 that such developments should be highly encouraged and adopted even in individual laboratories, as they might
543 promote individual projects and, in turn, in the field itself.

544
545

546 **References:**

- 547 1. Flemming, W. *Zellsubstanz, Kern und Zelltheilung*. F.C.W. Vogel, Leipzig (1882).
- 548 2. Heitz, E. Das Heterochromatin der Moose. *Jahrbücher für wissenschaftliche Bot.* (1928).
- 549 3. Boveri, T. *Die Blastomerenkerne von Ascaris megalocephala und die Theorie der Chromosomenindividualität*.
550 *Engelmann* (1909).
- 551 4. Cremer, T. *et al.* Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by
552 laser-UV-microirradiation experiments. *Hum. Genet.* **62**, 201–209 (1982).
- 553 5. Manuelidis, L. Individual interphase chromosome domains revealed by in situ hybridization. *Hum. Genet.* **71**,
554 288–293 (1985).
- 555 6. Schardin, M., Cremer, T., Hager, H. D. & Lang, M. Specific staining of human chromosomes in Chinese
556 hamster x man hybrid cell lines demonstrates interphase chromosome territories. *Hum. Genet.* **71**, 281–287
557 (1985).
- 558 7. Cremer, M. *et al.* Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. in
559 *The Nucleus. Methods in Molecular Biology* 205–239 (2012). doi:10.1007/978-1-59745-406-3_15
- 560 8. Branco, M. R. & Pombo, A. Intermingling of chromosome territories in interphase suggests role in
561 translocations and transcription-dependent associations. *PLoS Biol.* **4**, e138 (2006).
- 562 9. Chambeyron, S. & Bickmore, W. A. Chromatin decondensation and nuclear reorganization of the HoxB locus
563 upon induction of transcription. *Genes Dev.* **18**, 1119–1130 (2004).
- 564 10. Ferrai, C. *et al.* Poised transcription factories prime silent uPA gene prior to activation. *PLoS Biol.* **8**, e1000270
565 (2010).
- 566 11. Rosin, L. F., Nguyen, S. C. & Joyce, E. F. Condensin II drives large-scale folding and spatial partitioning of
567 interphase chromosomes in Drosophila nuclei. *PLOS Genet.* **14**, e1007393 (2018).
- 568 12. Su, J. H., Zheng, P., Kinrot, S. S., Bintu, B. & Zhuang, X. Genome-Scale Imaging of the 3D Organization and
569 Transcriptional Activity of Chromatin. *Cell* **182**, 1641–1659 (2020).
- 570 13. Fritz, A. J., Sehgal, N., Pliss, A., Xu, J. & Berezney, R. Chromosome territories and the global regulation of the
571 genome. *Genes, Chromosom. Cancer* **58**, 407–426 (2019).
- 572 14. Nguyen, H. Q. *et al.* 3D mapping and accelerated super-resolution imaging of the human genome using in
573 situ sequencing. *Nat. Methods* **17**, 822–832 (2020).
- 574 **First mention of OligoFISSEQ (DNA in situ sequencing); showcasing its multiplexing possibilities by imaging 249**
575 **loci simultaneously.**
- 576 15. Takei, Y. *et al.* Integrated spatial genomics reveals global architecture of single nuclei. *Nature* **590**, 344–350
577 (2021).
- 578 16. Payne, A. C. *et al.* In situ genome sequencing resolves DNA sequence and structure in intact biological
579 samples. *Science* **371**, eaay3446 (2020).
- 580 17. Cullen, K. E., Kladdé, M. P. & Seyfred, M. A. Interaction between transcription regulatory regions of prolactin
581 chromatin. *Science* **261**, 203–206 (1993).
- 582 18. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–
583 1311 (2002).
- 584 19. Simonis, M. *et al.* Nuclear organization of active and inactive chromatin domains uncovered by chromosome
585 conformation capture-on-chip (4C). *Nat Genet* **38**, 1348–1354 (2006).

- 586 20. Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for
587 mapping interactions between genomic elements. *Genome Res.* **16**, 1299–1309 (2006).
- 588 21. Schoenfelder, S. *et al.* The pluripotent regulatory circuitry connecting promoters to their long-range
589 interacting elements. *Genome Res.* **25**, 582–97 (2015).
- 590 22. Schoenfelder, S. *et al.* Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem
591 cell genome. *Nat Genet* **47**, 1179–1186 (2015).
- 592 23. Mifsud, B. *et al.* Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C.
593 *Nat Genet* **47**, 598–606 (2015).
- 594 24. Hughes, J. R. *et al.* Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-
595 throughput experiment. *Nat. Genet.* **46**, 205–212 (2014).
- 596 25. Davies, J. O. J. *et al.* Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat.*
597 *Methods* **13**, 74–80 (2015).
- 598 26. Jäger, R. *et al.* Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat. Commun.*
599 **6**, 6178 (2015).
- 600 27. Lieberman-Aiden, E. *et al.* Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of
601 the Human Genome. *Science* **326**, 289–293 (2009).
- 602 28. Duan, Z. *et al.* A three-dimensional model of the yeast genome. *Nature* **465**, 363–367 (2010).
- 603 29. Sati, S. & Cavalli, G. Chromosome conformation capture technologies and their impact in understanding
604 genome function. *Chromosoma* **126**, 33–44 (2017).
- 605 30. Horike, S. I., Cai, S., Miyano, M., Cheng, J. F. & Kohwi-Shigematsu, T. Loss of silent-chromatin looping and
606 impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.* **37**, 31–40 (2005).
- 607 31. Fullwood, M. J., Wei, C. L., Liu, E. T. & Ruan, Y. Next-generation DNA sequencing of paired-end tags (PET) for
608 transcriptome and genome analyses. *Genome Res.* **19**, 521–532 (2009).
- 609 32. Mumbach, M. R. *et al.* HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat.*
610 *Methods* **13**, 919–922 (2016).
- 611 33. Fang, R. *et al.* Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.*
612 **26**, 1345–1348 (2016).
- 613 34. Chen, Y. *et al.* Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a
614 cytological ruler. *J. Cell Biol.* **217**, 4025–4048 (2018).
- 615 **Introduction of TSA-seq, a cytological ruler for nuclear speckles and lamina. It is the first genomics method that**
616 **enabled transforming sequencing reads into physical distances in the nucleus.**
- 617 35. Quinodoz, S. A. *et al.* Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus.
618 *Cell* **174**, 744–757 (2018).
- 619 **Introduction of SPRITE and discovery of inter-chromosomal hubs around the nucleolus and nuclear speckles.**
- 620 36. Arrastia, M. V *et al.* A single-cell method to map higher-order 3D genome organization in thousands of
621 individual cells reveals structural heterogeneity in mouse ES cells. *Preprint at bioRxiv*
622 doi:10.1101/2020.08.11.242081 (2020).
- 623 37. Quinodoz, S. A. *et al.* RNA promotes the formation of spatial compartments in the nucleus. *Preprint at bioRxiv*
624 doi:10.1101/2020.08.25.267435 (2020).
- 625 38. Van Steensel, B. & Henikoff, S. Identification of in vivo DNA targets of chromatin proteins using tethered
626 Dam methyltransferase. *Nat. Biotechnol.* **14**, 424–428 (2000).
- 627 **Introduction of DamIG, which is used for spatial positioning of HP1 on chromatin in vitro and in vivo in Drosophila**
628 **melanogaster.**
- 629 39. Vogel, M. J., Peric-Hupkes, D. & van Steensel, B. Detection of in vivo protein - DNA interactions using DamID
630 in mammalian cells. *Nat. Protoc.* **2**, 1467–1478 (2007).
- 631 40. Guelen, L. *et al.* Domain organization of human chromosomes revealed by mapping of nuclear lamina
632 interactions. *Nature* **453**, 948–951 (2008).
- 633 41. Beagrie, R. A. *et al.* Complex multi-enhancer contacts captured by genome architecture mapping. *Nature*
634 **543**, 519–524 (2017).
- 635 **Introduction of the GAM method and discovery of a richness of multi-way contacts in the nucleus.**
- 636 42. Zhang, L. *et al.* TSA-seq reveals a largely conserved genome organization relative to nuclear speckles with
637 small position changes tightly correlated with gene expression changes. *Genome Res.* **31**, 251–264 (2021).
- 638 43. Szabo, Q. *et al.* TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci. Adv.*

- 639 4, eaar8082 (2018).
- 640 44. Bintu, B. *et al.* Super-resolution chromatin tracing reveals domains and cooperative interactions in single
641 cells. *Science* **362**, eaau1783 (2018).
- 642 **This study uses high-throughput oligopaint technology to study chromatin dynamics and shows that TADs are**
643 **highly stochastic in single cells, but predictable at the population level.**
- 644 45. Cardozo Gizzi, A. M. *et al.* Microscopy-Based Chromosome Conformation Capture Enables Simultaneous
645 Visualization of Genome Organization and Transcription in Intact Organisms. *Mol. Cell* **74**, 212-222.e5 (2019).
- 646 **Highly multiplexed sequential Oligo-FISH (RNA and DNA) based on the conjunction of microscopy and**
647 **microfluidics in *Drosophila melanogaster* embryos.**
- 648 46. Nir, G. *et al.* Walking along chromosomes with super-resolution imaging, contact maps, and integrative
649 modeling. *PLoS Genet.* **14**, e1007872 (2018).
- 650 47. Mateo, L. J. *et al.* Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54
651 (2019).
- 652 48. Sawah, A. N. *et al.* Lamina-Dependent Stretching and Unconventional Chromosome Compartments in Early *C.*
653 *elegans* Embryos. *Mol. Cell* **78**, 96–111 (2020).
- 654 49. Gu, B. *et al.* Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements.
655 *Science* **359**, 1050–1055 (2018).
- 656 50. Shaban, H. A. & Seeber, A. Monitoring the spatio-temporal organization and dynamics of the genome.
657 *Nucleic Acids Res.* **48**, 3423–3434 (2020).
- 658 51. Barth, R., Bystricky, K. & Shaban, H. A. Coupling chromatin structure and dynamics by live super-resolution
659 imaging. *Sci. Adv.* **6**, eaaz2196 (2020).
- 660 52. Vangala, P. *et al.* High-Resolution Mapping of Multiway Enhancer-Promoter Interactions Regulating
661 Pathogen Detection. *Mol. Cell* **80**, 359–373 (2020).
- 662 53. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**,
663 381–385 (2012).
- 664 54. Sexton, T. *et al.* Three-dimensional folding and functional organization principles of the *Drosophila* genome.
665 *Cell* **148**, 458–472 (2012).
- 666 55. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin
667 interactions. *Nature* **485**, 376–380 (2012).
- 668 56. Kalhor, R. *et al.* Developmental barcoding of whole mouse via homing CRISPR. *Science* **361**, eaat9804 (2018).
- 669 57. Szabo, Q. *et al.* Regulation of single-cell genome organization into TADs and chromatin nanodomains. *Nat.*
670 *Genet.* **52**, 1151–1157 (2020).
- 671 **First description of chromatin nanodomains in mammals using Oligopaints in conjunction with SIM.**
- 672 58. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* **15**, 2038–49 (2016).
- 673 59. Wutz, G. *et al.* Topologically associating domains and chromatin loops depend on cohesin and are regulated
674 by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **36**, 3573–3599 (2017).
- 675 60. Rao, S. S. P. *et al.* Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305–320 (2017).
- 676 61. Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature*
677 **551**, 51–56 (2017).
- 678 62. Guo, Y. *et al.* CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*
679 **162**, 900–910 (2015).
- 680 63. Haarhuis, J. H. I. *et al.* The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**, 693–
681 707 (2017).
- 682 64. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from
683 Genomic Compartmentalization. *Cell* **169**, 930–944 (2017).
- 684 65. Bonev, B. *et al.* Multiscale 3D Genome Rewiring during Mouse Article Multiscale 3D Genome Rewiring during
685 Mouse Neural Development. *Cell* **171**, 557–572 (2017).
- 686 66. Ulianov, S. V. *et al.* Active chromatin and transcription play a key role in chromosome partitioning into
687 topologically associating domains. *Genome Res.* **26**, 70–84 (2016).
- 688 67. Sanborn, A. L. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type
689 and engineered genomes. *Proc Natl Acad Sci U S A* **112**, 6456–6465 (2015).
- 690 68. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay
691 of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6697–6706 (2018).

- 692 69. Rowley, M. J. *et al.* Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Mol. Cell* **67**, 837–
693 852 (2017).
- 694 70. Benedetti, F., Racko, D., Dorier, J., Burnier, Y. & Stasiak, A. Transcription-induced supercoiling explains
695 formation of self-interacting chromatin domains in *S. Pombe*. *Nucleic Acids Res.* **45**, 9850–9859 (2017).
- 696 71. Szabo, Q., Bantignies, F. & Cavalli, G. Principles of genome folding into topologically associating domains. *Sci.*
697 *Adv.* **5**, eaaw1668 (2019).
- 698 72. Jerković, I., Szabo, Q., Bantignies, F. & Cavalli, G. Higher-Order Chromosomal Structures Mediate Genome
699 Function. *J. Mol. Biol.* **432**, 676–681 (2020).
- 700 73. Oomen, M. E., Hedger, A. K., Watts, J. K. & Dekker, J. Detecting chromatin interactions between and along
701 sister chromatids with SisterC. *Nat. Methods* **17**, 1002–1009 (2020).
- 702 74. Mitter, M. *et al.* Conformation of sister chromatids in the replicated human genome. *Nature* **586**, 139–144
703 (2020).
- 704 75. AlHajj Abed, J. *et al.* Highly structured homolog pairing reflects functional organization of the *Drosophila*
705 genome. *Nat. Commun.* **10**, 1–14 (2019).
- 706 76. Franke, M. *et al.* Formation of new chromatin domains determines pathogenicity of genomic duplications.
707 *Nature* **538**, 265–269 (2016).
- 708 77. Lupiáñez, D. G., Spielmann, M. & Mundlos, S. Breaking TADs: How Alterations of Chromatin Domains Result
709 in Disease. *Trends Genet.* **32**, 225–237 (2016).
- 710 78. Despang, A. *et al.* Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles
711 of TAD architecture. *Nat. Genet.* **51**, 1263–1271 (2019).
- 712 79. Andrey, G. & Mundlos, S. The three-dimensional genome: regulating gene expression during pluripotency
713 and development. *Development* **144**, 3646–3658 (2017).
- 714 80. Weischenfeldt, J. *et al.* Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in
715 enhancer hijacking. *Nat. Genet.* **49**, 65–74 (2016).
- 716 81. Liu, X. S. *et al.* Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. *Cell*
717 **172**, 979–992 (2018).
- 718 82. Bruneau, B. G. & Nora, E. P. Chromatin Domains Go on Repeat in Disease. *Cell* **175**, 38–40 (2018).
- 719 83. Akdemir, K. C. *et al.* Somatic mutation distributions in cancer genomes vary with three-dimensional
720 chromatin structure. *Nat. Genet.* **52**, 1178–1188 (2020).
- 721 84. Valton, A. L. & Dekker, J. TAD disruption as oncogenic driver. *Curr. Opin. Genet. Dev.* **36**, 34–40 (2016).
- 722 85. Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin
723 looping. *Cell* **159**, 1665–1680 (2014).
- 724 **This paper showcased the first use of the *in situ* Hi-C protocol, and introduced HICCUPS, a contacts caller dedicated**
725 **to chromatin loops.**
- 726 86. Nagano, T. *et al.* Comparison of Hi-C results using in-solution versus in-nucleus ligation. *Genome Biol.* **16**, 175
727 (2015).
- 728 87. Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64
729 (2013).
- 730 88. Ma, W. *et al.* Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA
731 genes. *Nat. Methods* **12**, 71–78 (2014).
- 732 89. Ramani, V. *et al.* Mapping 3D genome architecture through in situ DNase Hi-C. *Nat. Protoc.* **11**, 2104–2121
733 (2016).
- 734 90. Deng, X. *et al.* Bipartite structure of the inactive mouse X chromosome. *Genome Biol.* **16**, 152 (2015).
- 735 91. Hsieh, T. H. S. *et al.* Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* **162**,
736 108–119 (2015).
- 737 92. Hsieh, T.-H. S. *et al.* Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Mol.*
738 *Cell* **78**, 539–553 (2020).
- 739 **High resolution Micro-C performed in mESC with concomitant analysis of nucleosome positioning.**
- 740 93. Krietenstein, N. *et al.* Ultrastructural Details of Mammalian Chromosome Architecture. *Mol. Cell* **78**, 554–
741 565 (2020).
- 742 **High resolution Micro-C performed in human ESC and fibroblasts with concomitant analysis of nucleosome**
743 **positioning.**
- 744 94. Goel, V. Y. & Hansen, A. S. The macro and micro of chromosome conformation capture. *WIREs Dev. Biol.*

745 e395 (2020). doi:10.1002/wdev.395

746 95. Akgol Oksuz, B. *et al.* Systematic evaluation of chromosome conformation capture assays. *Preprint at bioRxiv*
747 doi:10.1101/2020.12.26.424448 (2020).

748 96. Baranello, L., Kouzine, F., Sanford, S. & Levens, D. ChIP bias as a function of cross-linking time. *Chromosom.*
749 *Res.* **24**, 175–181 (2016).

750 97. Gavrilov, A., Razin, S. V. & Cavalli, G. In vivo formaldehyde cross-linking: It is time for black box analysis. *Brief.*
751 *Funct. Genomics* **14**, 163–165 (2015).

752 98. Kempfer, R. & Pombo, A. Methods for mapping 3D chromosome architecture. *Nat. Rev. Genet.* **21**, 207–226
753 (2020).

754 99. Fullwood, M. J. & Ruan, Y. ChIP-based methods for the identification of long-range chromatin interactions.
755 *J. Cell. Biochem.* **107**, 30–39 (2009).

756 100. Kind, J. *et al.* Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. *Cell* **163**, 134–147
757 (2015).

758 101. Redolfi, J. *et al.* DamC reveals principles of chromatin folding in vivo without crosslinking and ligation. *Nat.*
759 *Struct. Mol. Biol.* **26**, 471–480 (2019).

760 102. Li, X. *et al.* Long-read ChIA-PET for base-pair-resolution mapping of haplotype-specific chromatin interactions.
761 *Nat. Protoc.* **12**, 899–915 (2017).

762 103. Andrey, G. *et al.* Characterization of hundreds of regulatory landscapes in developing limbs reveals two
763 regimes of chromatin folding. *Genome Res.* **27**, 223–233 (2017).

764 104. Rooijers, K. *et al.* Simultaneous quantification of protein–DNA contacts and transcriptomes in single cells.
765 *Nat. Biotechnol.* **37**, 766–772 (2019).

766 105. Sahl, S. J., Hell, S. W. & Jakobs, S. Fluorescence nanoscopy in cell biology. *Nature Reviews Molecular Cell*
767 *Biology* **18**, 685–701 (2017).

768 106. Liu, Z., Lavis, L. D. & Betzig, E. Imaging Live-Cell Dynamics and Structure at the Single-Molecule Level.
769 *Molecular Cell* **58**, 644–659 (2015).

770 107. Heintzmann, R. & Cremer, C. G. Laterally modulated excitation microscopy: improvement of resolution by
771 using a diffraction grating. in *Optical Biopsies and Microscopic Techniques III* **3568**, (SPIE, 1999).

772 108. Gustafsson, M. G. L. Surpassing the lateral resolution limit by a factor of two using structured illumination
773 microscopy. *J. Microsc.* **198**, 82–87 (2000).

774 109. Gustafsson, M. G. L. *et al.* Three-dimensional resolution doubling in wide-field fluorescence microscopy by
775 structured illumination. *Biophys. J.* **94**, 4957–4970 (2008).

776 110. Schermelleh, L. *et al.* Subdiffraction multicolor imaging of the nuclear periphery with 3D structured
777 illumination microscopy. *Science* **320**, 1332–1336 (2008).

778 111. Miron, E. *et al.* Chromatin arranges in chains of mesoscale domains with nanoscale functional topography
779 independent of cohesin. *Sci. Adv.* **6**, eaba8811 (2020).

780 112. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction
781 microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006).

782 113. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645
783 (2006).

784 114. Hess, S. T., Girirajan, T. P. K. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation
785 localization microscopy. *Biophys. J.* **91**, 4258–4272 (2006).

786 115. Huang, B., Babcock, H. & Zhuang, X. Breaking the diffraction barrier: Super-resolution imaging of cells. *Cell*
787 **143**, 1047–1058 (2010).

788 116. Ricci, M. A., Manzo, C., Lakadamyali, M. & Cosma, M. P. Chromatin Fibers Are Formed by Heterogeneous
789 Groups of Nucleosomes In Vivo. *Cell* **160**, 1145–1158 (2015).

790 117. Nozaki, T. *et al.* Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging.
791 *Mol. Cell* **67**, 282–293 (2017).

792 118. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-
793 emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).

794 119. Klar, T. A., Jakobs, S., Dyba, M., Egner, A. & Hell, S. W. Fluorescence microscopy with diffraction resolution
795 barrier broken by stimulated emission. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8206–8210 (2000).

796 120. Hofmann, M., Eggeling, C., Jakobs, S. & Hell, S. W. Breaking the diffraction barrier in fluorescence microscopy
797 at low light intensities by using reversibly photoswitchable proteins. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17565–

- 798 17569 (2005).
- 799 121. Gwosch, C. K. *et al.* MINFLUX nanoscopy delivers multicolor nanometer 3D-resolution in (living) cells. *Nat. Methods* **17**, 217–224 (2020).
- 800
- 801 122. Beliveau, B. J. *et al.* Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 21301–21306 (2012).
- 802
- 803 123. Wang, S. *et al.* Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016).
- 804
- 805 124. Ou, H. D. *et al.* ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* **357**, eaag0025 (2017).
- 806
- 807 125. Yaffe, E. & Tanay, A. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nat. Genet.* **43**, 1059–1065 (2011).
- 808
- 809 126. Imakaev, M. *et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **9**, 999–1003 (2012).
- 810
- 811 127. Ryba, T. *et al.* Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* **20**, 761–770 (2010).
- 812
- 813 128. Di Pierro, M., Cheng, R. R., Aiden, E. L., Wolynes, P. G. & Onuchic, J. N. De novo prediction of human chromosome structures: Epigenetic marking patterns encode genome architecture. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12126–12131 (2017).
- 814
- 815 129. Lin, D., Bonora, G., Yardimci, G. G. & Noble, W. S. Computational methods for analyzing and modeling genome structure and organization. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **11**, e1435 (2019).
- 816
- 817 130. 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 (2012).
- 818
- 819 131. Crane, E. *et al.* Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* **523**, 240–244 (2015).
- 820
- 821 132. Shin, H. *et al.* TopDom: An efficient and deterministic method for identifying topological domains in genomes. *Nucleic Acids Res.* **44**, e70 (2015).
- 822
- 823 **One of the most reliable and popular TAD callers.**
- 824
- 825 133. Malik, L. & Patro, R. Rich chromatin structure prediction from Hi-C data. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **16**, 1448–1458 (2019).
- 826
- 827 134. Haddad, N., Vaillant, C. & Jost, D. IC-Finder: Inferring robustly the hierarchical organization of chromatin folding. *Nucleic Acids Res.* **45**, 81 (2017).
- 828
- 829 135. Soler-Vila, P., Cuscó, P., Farabella, I., Di Stefano, M. & Marti-Renom, M. A. Hierarchical chromatin organization detected by TADpole. *Nucleic Acids Res.* **48**, e39 (2020).
- 830
- 831 **One of the most recently developed TAD callers; this paper is extremely valuable for its benchmarking quality and for multiple TAD caller comparisons.**
- 832
- 833 136. Norton, H. K. *et al.* Detecting hierarchical genome folding with network modularity. *Nat. Methods* **15**, 119–122 (2018).
- 834
- 835 137. Moller, J. & de Pablo, J. J. Bottom-Up Meets Top-Down: The Crossroads of Multiscale Chromatin Modeling. *Biophysical Journal* **118**, 2057–2065 (2020).
- 836
- 837 138. Bendandi, A., Dante, S., Zia, S. R., Diaspro, A. & Rocchia, W. Chromatin Compaction Multiscale Modeling: A Complex Synergy Between Theory, Simulation, and Experiment. *Frontiers in Molecular Biosciences* **7**, 15 (2020).
- 838
- 839
- 840 139. Zufferey, M., Tavernari, D., Oricchio, E. & Ciriello, G. Comparison of computational methods for the identification of topologically associating domains. *Genome Biol.* **19**, 217 (2018).
- 841
- 842 140. Forcato, M. *et al.* Comparison of computational methods for Hi-C data analysis. *Nat. Methods* **14**, 679–685 (2017).
- 843
- 844 141. Ay, F., Bailey, T. L. & Noble, W. S. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res.* **24**, 999–1011 (2014).
- 845
- 846 142. Carty, M. *et al.* An integrated model for detecting significant chromatin interactions from high-resolution Hi-C data. *Nat. Commun.* **8**, 1–10 (2017).
- 847
- 848 143. Pal, K., Forcato, M. & Ferrari, F. Hi-C analysis: from data generation to integration. *Biophysical Reviews* **11**, 67–78 (2019).
- 849
- 850 144. Dozmorov, M., Sirusb; & Benfeitas, R. Hi-C data analysis tools and papers. Available at:

851 https://github.com/mdozmorov/HiC_tools/blob/master/README.md.

852 145. Roayaei Ardakany, A., Gezer, H. T., Lonardi, S. & Ay, F. Mustache: Multi-scale detection of chromatin loops
853 from Hi-C and Micro-C maps using scale-space representation. *Genome Biol.* **21**, 256 (2020).

854 146. Jordan Rowley, M. *et al.* Analysis of Hi-C data using SIP effectively identifies loops in organisms from C.
855 *Elegans* to mammals. *Genome Res.* **31**, (2020).

856 147. Ay, F. & Noble, W. S. Analysis methods for studying the 3D architecture of the genome. *Genome Biology* **16**,
857 183 (2015).

858 148. Nikumbh, S. & Pfeifer, N. Genetic sequence-based prediction of long-range chromatin interactions suggests
859 a potential role of short tandem repeat sequences in genome organization. *BMC Bioinformatics* **18**, 218
860 (2017).

861 149. Rhie, S. K. *et al.* Using 3D epigenomic maps of primary olfactory neuronal cells from living individuals to
862 understand gene regulation. *Sci. Adv.* **4**, eaav8550 (2018).

863 150. Hafez, D. *et al.* McEnhancer: Predicting gene expression via semi-supervised assignment of enhancers to
864 target genes. *Genome Biol.* **18**, 199 (2017).

865 151. Di Iulio, J. *et al.* The human noncoding genome defined by genetic diversity. *Nat. Genet.* **50**, 333–337 (2018).

866 152. Javierre, B. M. *et al.* Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants
867 to Target Gene Promoters. *Cell* **167**, 1369–1384 (2016).

868 153. Mitchelmore, J., Grinberg, N. F., Wallace, C. & Spivakov, M. Functional effects of variation in transcription
869 factor binding highlight long-range gene regulation by epromoters. *Nucleic Acids Res.* **48**, 2866–2879 (2020).

870 154. Malysheva, V., Mendoza-Parra, M. A., Blum, M., Spivakov, M. & Gronemeyer, H. Gene regulatory network
871 reconstruction incorporating 3D chromosomal architecture reveals key transcription factors and DNA
872 elements driving neural lineage commitment. *Preprint at bioRxiv* doi:10.1101/303842 (2019).

873 155. Madrid-Mencia, M., Raineri, E., Cao, T. B. N. & Pancaldi, V. Using GARDEN-NET and ChAseR to explore human
874 haematopoietic 3D chromatin interaction networks. *Nucleic Acids Res.* **48**, 4066–4080 (2020).

875 156. Zhang, Y. *et al.* Enhancing Hi-C data resolution with deep convolutional neural network HiCPlus. *Nat.*
876 *Commun.* **9**, 1–9 (2018).

877 157. Oluwadare, O., Highsmith, M. & Cheng, J. An Overview of Methods for Reconstructing 3-D Chromosome and
878 Genome Structures from Hi-C Data. *Biological Procedures Online* **21**, 1–20 (2019).

879 158. Marti-Renom, M. A. *et al.* Challenges and guidelines toward 4D nucleome data and model standards. *Nature*
880 *Genetics* **50**, 1352–1358 (2018).

881 159. MacPherson, Q., Beltran, B. & Spakowitz, A. J. Bottom-up modeling of chromatin segregation due to
882 epigenetic modifications. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 12739–12744 (2018).

883 160. Bianco, S., Chiariello, A. M., Annunziatella, C., Esposito, A. & Nicodemi, M. Predicting chromatin architecture
884 from models of polymer physics. *Chromosome Research* **25**, 25–34 (2017).

885 161. Mirny, L. A. The fractal globule as a model of chromatin architecture in the cell. *Chromosom. Res.* **19**, 37–51
886 (2011).

887 162. Barbieri, M. *et al.* A model of the large-scale organization of chromatin. *Biochemical Society Transactions* **41**,
888 508–512 (2013).

889 163. Jost, D., Carrivain, P., Cavalli, G. & Vaillant, C. Modeling epigenome folding: Formation and dynamics of
890 topologically associated chromatin domains. *Nucleic Acids Res.* **42**, 9553–9561 (2014).

891 164. Jost, D. & Vaillant, C. Epigenomics in 3D: Importance of long-range spreading and specific interactions in
892 epigenomic maintenance. *Nucleic Acids Res.* **46**, 2252–2264 (2018).

893 165. Barbieri, M. *et al.* Active and poised promoter states drive folding of the extended HoxB locus in mouse
894 embryonic stem cells. *Nat. Struct. Mol. Biol.* **24**, 515–524 (2017).

895 166. Falk, M. *et al.* Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature*
896 **570**, 395–399 (2019).

897 167. Oliveira Junior, A. B., Contessoto, V. G., Mello, M. F. & Onuchic, J. N. A Scalable Computational Approach for
898 Simulating Complexes of Multiple Chromosomes. *J. Mol. Biol.* (2020). doi:10.1016/j.jmb.2020.10.034

899 168. Ghosh, S. K. & Jost, D. How epigenome drives chromatin folding and dynamics, insights from efficient coarse-
900 grained models of chromosomes. *PLoS Comput. Biol.* **14**, e1006159 (2018).

901 169. Giorgetti, L. *et al.* Predictive polymer modeling reveals coupled fluctuations in chromosome conformation
902 and transcription. *Cell* **157**, 950–963 (2014).

903 170. Bianco, S. *et al.* Polymer physics predicts the effects of structural variants on chromatin architecture. *Nat.*

- 904 *Genet.* **50**, 662–667 (2018).
- 905 171. Contessoto, V. G. *et al.* The Nucleome Data Bank: web-based resources to simulate and analyze the three-
906 dimensional genome. *Nucleic Acids Res.* **49**, 172–182 (2020).
- 907 172. Rousseau, M., Fraser, J., Ferraiuolo, M. A., Dostie, J. & Blanchette, M. Three-dimensional modeling of
908 chromatin structure from interaction frequency data using Markov chain Monte Carlo sampling. *BMC*
909 *Bioinformatics* **12**, 414 (2011).
- 910 173. Baú, D. *et al.* The three-dimensional folding of the α -globin gene domain reveals formation of chromatin
911 globules. *Nat. Struct. Mol. Biol.* **18**, 107–115 (2011).
- 912 174. Trussart, M. *et al.* Defined chromosome structure in the genome-reduced bacterium *Mycoplasma*
913 *pneumoniae*. *Nat. Commun.* **8**, 14665 (2017).
- 914 175. Zhu, G. *et al.* Reconstructing spatial organizations of chromosomes through manifold learning. *Nucleic Acids*
915 *Res.* **46**, e50 (2018).
- 916 176. Paulsen, J. *et al.* Long-range interactions between topologically associating domains shape the four-
917 dimensional genome during differentiation. *Nat. Genet.* **51**, 835–843 (2019).
- 918 177. Lesne, A., Riposo, J., Roger, P., Cournac, A. & Mozziconacci, J. 3D genome reconstruction from chromosomal
919 contacts. *Nat. Methods* **11**, 1141–1143 (2014).
- 920 178. Trieu, T., Oluwadare, O. & Cheng, J. Hierarchical Reconstruction of High-Resolution 3D Models of Large
921 Chromosomes. *Sci. Rep.* **9**, 1–12 (2019).
- 922 179. Tan, L., Xing, D., Chang, C.-H. H., Li, H. & Xie, X. S. Three-dimensional genome structures of single diploid
923 human cells. *Science* **361**, 924–928 (2018).
- 924 180. Di Stefano, M. *et al.* Transcriptional activation during cell reprogramming correlates with the formation of
925 3D open chromatin hubs. *Nat. Commun.* **11**, 1–12 (2020).
- 926 181. Li, Q. *et al.* The three-dimensional genome organization of *Drosophila melanogaster* through data integration.
927 *Genome Biol.* **18**, 145 (2017).
- 928 182. Hua, N. *et al.* Producing genome structure populations with the dynamic and automated PGS software. *Nat.*
929 *Protoc.* **13**, 915–926 (2018).
- 930 183. Whalen, S., Truty, R. M. & Pollard, K. S. Enhancer-promoter interactions are encoded by complex genomic
931 signatures on looping chromatin. *Nat. Genet.* **48**, 488–496 (2016).
- 932 184. Li, W., Wong, W. H. & Jiang, R. DeepTACT: Predicting 3D chromatin contacts via bootstrapping deep learning.
933 *Nucleic Acids Res.* **47**, e60 (2019).
- 934 185. Fudenberg, G., Kelley, D. R. & Pollard, K. S. Predicting 3D genome folding from DNA sequence with Akita. *Nat.*
935 *Methods* **17**, 1111–1117 (2020).
- 936 **First computational modelling paper describing modelling of chromatin conformation solely from the DNA**
- 937 186. Schwesinger, R. *et al.* DeepC: predicting 3D genome folding using megabase-scale transfer learning. *Nat.*
938 *Methods* **17**, 1118–1124 (2020).
- 939 **One of the first computational tools to model chromatin conformation solely from the DNA sequence. However,**
940 **unlike Akita, it requires training on the epigenetics data.**
- 941 187. Zhang, S., Chasman, D., Knaack, S. & Roy, S. In silico prediction of high-resolution Hi-C interaction matrices.
942 *Nat. Commun.* **10**, 1–18 (2019).
- 943 188. Belokopytova, P. S., Nuriddinov, M. A., Mozheiko, E. A., Fishman, D. & Fishman, V. Quantitative prediction
944 of enhancer–promoter interactions. *Genome Res.* **30**, 72–84 (2020).
- 945 189. Qi, Y. *et al.* Data-driven Polymer Model for Mechanistic Exploration of Diploid Genome Organization. *Biophys.*
946 *J.* **119**, 1905–1916 (2020).
- 947 190. de Luca, K. L. & Kind, J. Single-cell damid to capture contacts between dna and the nuclear lamina in
948 individual mammalian cells. in *Methods in Molecular Biology* **2157**, 159–172 (Humana Press Inc., 2021).
- 949 191. Li, G. *et al.* Joint profiling of DNA methylation and chromatin architecture in single cells. *Nat. Methods* **16**,
950 991–993 (2019).
- 951 192. Zheng, M. *et al.* Multiplex chromatin interactions with single-molecule precision. *Nature* **566**, 558–562
952 (2019).
- 953 193. Ramani, V. *et al.* Massively multiplex single-cell Hi-C. *Nat. Methods* **14**, 263–266 (2017).
- 954 194. You, Q. *et al.* Direct DNA crosslinking with CAP-C uncovers transcription-dependent chromatin organization
955 at high resolution. *Nat. Biotechnol.* **39**, 225–235 (2020).
- 956 195. Girelli, G. *et al.* GPSeq reveals the radial organization of chromatin in the cell nucleus. *Nat. Biotechnol.* **38**,

- 957 1184–1193 (2020).
- 958 196. Wang, Y. *et al.* SPIN reveals genome-wide landscape of nuclear compartmentalization. *Genome Biol.* **22**, 36
959 (2021).
- 960 **First tool that combines nuclear architecture (positioning) data from TSA-seq and DamID in order to compute**
961 **spatial compartmentalization of chromatin in respect to the lamina or nuclear bodies such as nuclear**
962 **speckles.**
- 963 197. Kubalová, I. *et al.* Prospects and limitations of expansion microscopy in chromatin ultrastructure
964 determination. *Chromosom. Res.* **28**, 355–368 (2020).
- 965 198. Shi, X. *et al.* Label-retention expansion microscopy. *Preprint at bioRxiv* doi:10.1101/687954 (2021).
- 966 199. Wassie, A. T., Zhao, Y. & Boyden, E. S. Expansion microscopy: principles and uses in biological research.
967 *Nature Methods* **16**, 33–41 (2019).
- 968 200. Alon, S. *et al.* Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems.
969 *Science* **371**, eaax2656 (2021).
- 970 201. Robinett, C. C. *et al.* In vivo localization of DNA sequences and visualization of large-scale chromatin
971 organization using lac operator/repressor recognition. *J. Cell Biol.* **135**, 1685–1700 (1996).
- 972 202. Belmont, A. S. & Straight, A. F. In vivo visualization of chromosomes using lac operator-repressor binding.
973 *Trends Cell Biol.* **8**, 121–124 (1998).
- 974 203. Lucas, J. S., Zhang, Y., Dudko, O. K. & Murre, C. 3D trajectories adopted by coding and regulatory DNA
975 elements: First-passage times for genomic interactions. *Cell* **158**, 339–352 (2014).
- 976 204. Chen, H. *et al.* Dynamic interplay between enhancer–promoter topology and gene activity. *Nat. Genet.* **50**,
977 1296–1303 (2018).
- 978 205. Maass, P. G. *et al.* Spatiotemporal allele organization by allele-specific CRISPR live-cell imaging (SNP-CLING).
979 *Nat. Struct. Mol. Biol.* **25**, 176–184 (2018).
- 980 206. Wang, H. *et al.* CRISPR-mediated live imaging of genome editing and transcription. *Science* **365**, 1301–1305
981 (2019).
- 982 207. Geng, Y. & Pertsinidis, A. CAS-LiveFISH: Simple and versatile imaging of genomic loci in live mammalian cells
983 and early pre-implantation embryos. *Preprint at bioRxiv* doi:10.1101/2020.08.25.265306 (2020).
- 984 208. Neguembor, M. V. *et al.* (Po)STAC (Polycistronic SunTag modified CRISPR) enables live-cell and fixed-cell
985 super-resolution imaging of multiple genes. *Nucleic Acids Res.* **46**, e30 (2018).
- 986 209. Chu, L. *et al.* The 3D Topography of Mitotic Chromosomes. *Mol. Cell* **79**, 902–916.e6 (2020).
- 987 210. Liu, X. *et al.* In Situ Capture of Chromatin Interactions by Biotinylated dCas9. *Cell* **170**, 1028–1043 (2017).
- 988 211. Tsui, C. *et al.* dCas9-targeted locus-specific protein isolation method identifies histone gene regulators. *Proc.*
989 *Natl. Acad. Sci. U. S. A.* **115**, 2734–2741 (2018).
- 990 212. Myers, S. A. *et al.* Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-
991 mediated proximity labeling. *Nat. Methods* **15**, 437–439 (2018).
- 992 213. de Wit, E. *et al.* CTCF Binding Polarity Determines Chromatin Looping. *Mol. Cell* **60**, 676–684 (2015).
- 993 214. Gómez-Marín, C. *et al.* Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral
994 topological associating domains borders. *Proc Natl Acad Sci U S A* **112**, 7542–7547 (2015).
- 995 215. Bantignies, F. *et al.* Polycomb-dependent regulatory contacts between distant hox loci in drosophila. *Cell*
996 **144**, 214–226 (2011).
- 997 216. Monahan, K., Horta, A. & Lomvardas, S. LHX2- and LDB1-mediated trans interactions regulate olfactory
998 receptor choice. *Nature* **565**, 448–453 (2019).
- 999 217. Phillips-Cremins, J. E. *et al.* Architectural protein subclasses shape 3D organization of genomes during lineage
1000 commitment. *Cell* **153**, 1281–1295 (2013).
- 1001 218. Bhattacharyya, S., Chandra, V., Vijayanand, P. & Ay, F. Identification of significant chromatin contacts from
1002 HiChIP data by FitHiChIP. *Nat. Commun.* **10**, 1–14 (2019).
- 1003 219. Takei, Y. *et al.* Global architecture of the nucleus in single cells by DNA seqFISH+ and multiplexed
1004 immunofluorescence. *Preprint at bioRxiv* doi:10.1101/2020.11.29.403055 (2020).
- 1005 220. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between
1006 hypersensitive sites in the active beta-globin locus. *Mol Cell* **10**, 1453–1465 (2002).
- 1007 221. Palstra, R. J. *et al.* The β -globin nuclear compartment in development and erythroid differentiation. *Nat.*
1008 *Genet.* **35**, 190–194 (2003).
- 1009 222. Splinter, E. *et al.* CTCF mediates long-range chromatin looping and local histone modification in the β -globin

- 1010 locus. *Genes Dev.* **20**, 2349–2354 (2006).
- 1011 223. Vakoc, C. R. *et al.* Proximity among distant regulatory elements at the β -globin locus requires GATA-1 and
1012 FOG-1. *Mol. Cell* **17**, 453–462 (2005).
- 1013 224. Splinter, E. & De Laat, W. The complex transcription regulatory landscape of our genome: Control in three
1014 dimensions. *EMBO J.* **30**, 4345–4355 (2011).
- 1015 225. Zhao, Z. *et al.* Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically
1016 regulated intra- and interchromosomal interactions. *Nat. Genet.* **38**, 1341–1347 (2006).
- 1017 226. Andrey, G. *et al.* A switch between topological domains underlies HoxD genes collinearity in mouse limbs.
1018 *Science* **340**, 1234167 (2013).
- 1019 227. Ghavi-Helm, Y. *et al.* Enhancer loops appear stable during development and are associated with paused
1020 polymerase. *Nature* **512**, 96–100 (2014).
- 1021 228. Apostolou, E. *et al.* Genome-wide chromatin interactions of the nanog locus in pluripotency, differentiation,
1022 and reprogramming. *Cell Stem Cell* **12**, 699–712 (2013).
- 1023 229. Van De Werken, H. J. G. *et al.* 4C technology: Protocols and data analysis. in *Methods in Enzymology* **513**,
1024 89–112 (Academic Press Inc., 2012).
- 1025 230. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
- 1026 231. Y, G.-H. *et al.* Highly rearranged chromosomes reveal uncoupling between genome topology and gene
1027 expression. *Nat. Genet.* **51**, 1272–1282 (2019).
- 1028 232. Lupiáñez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-
1029 enhancer interactions. *Cell* **161**, 1012–1025 (2015).
- 1030 233. Ogiyama, Y., Schuettengruber, B., Papadopoulos, G. L., Chang, J.-M. & Cavalli, G. Polycomb-Dependent
1031 Chromatin Looping Contributes to Gene Silencing during Drosophila Development. *Mol. Cell* **71**, 73–88
1032 (2018).
- 1033 234. Flavahan, W. A. *et al.* Altered chromosomal topology drives oncogenic programs in SDH-deficient GISTs.
1034 *Nature* **575**, 229–233 (2019).
- 1035 235. Tarjan, D. R., Flavahan, W. A. & Bernstein, B. E. Epigenome editing strategies for the functional annotation
1036 of CTCF insulators. *Nat. Commun.* **10**, 4258 (2019).
- 1037 236. Kim, J. H. *et al.* LADL: light-activated dynamic looping for endogenous gene expression control. *Nat. Methods*
1038 **16**, 633–639 (2019).
- 1039 237. Wang, H. *et al.* CRISPR-Mediated Programmable 3D Genome Positioning and Nuclear Organization. *Cell* **175**,
1040 1405–1417 (2018).

1041
1042
1043

1044 **Acknowledgements**

1045 We thank Marco Di-Stefano for critical reading of the computational approaches section. We thank Bernd
1046 Schuttengrueber and Frederic Bantignies for input on figures. We thank Quentin Szabo for help with and
1047 input on Figure 4. I.J. was supported by EMBO Long-Term Fellowship (ALTF 559-2018) and the Laboratory
1048 of Excellence EpiGenMed. Research in the G.C. laboratory is supported by grants from the European
1049 Research Council (Advanced Grant 3DEpi, under grant agreement No 788972), the European Union’s
1050 Horizon 2020 research and innovation programme (MuG, under grant agreement No 676556 and
1051 ChromDesign, under the Martie Sklodowska-Curie grant agreement No 813327), the Agence Nationale de
1052 la Recherche (ANR-15-CE12-0006 EpiDevoMath), the Fondation pour la Recherche Médicale

1053 (DEI20151234396), the MSDAVENIR foundation (Project GENE-IGH), the INSERM and the French National
1054 Cancer Institute (INCa).

1055

1056 **Author contributions**

1057 The authors contributed equally to all aspects of the article.

1058 **Competing interests**

1059 The authors declare no competing interests.

1060 **Peer review information**

1061 *Nature Reviews Molecular Cell Biology* thanks Zhijun Duan; Bing Ren; Juan Vaquerizas, who co-reviewed
1062 with Elizabeth Ing-Simmons; and the other, anonymous reviewer(s) for their contribution to the peer
1063 review of this work.

1064 **Publisher's note**

1065 Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional
1066 affiliations.

1067

1068

1069 **Figure legends:**

1070 **Figure 1 Multiple levels of genome organization and the methods to study them.**

1071 DNA folds at multiple scales (indicated in the left) to build chromosomes. DNA winding around histones
1072 forms nucleosomes, which are organized into clutches, each containing ~1–2 kb of DNA¹¹⁶. Nucleosome
1073 clutches form chromatin nanodomains (CNDs) of ~100 kb in size, where most enhancer–promoter (E–P)
1074 contacts take place^{43,57,117}. At the scale of ~1 Mb, CNDs and CTCF–cohesin-dependent chromatin loops
1075 form topologically associating domains (TADs)^{43,57,58,67,116,213,214}. On the higher scale of up to 100s of Mbs,
1076 chromatin segregates into gene active and inactive compartments (A and B, respectively) and into
1077 compartment-specific contact hubs (not shown). At the highest topological level, the nucleus is organised
1078 into chromosome territories^{27,35}. Different techniques can be used to study different genome organization
1079 levels, and some techniques can be used to study different organization levels. Although Hi-C, genome
1080 architecture mapping (GAM) and other sequencing-based techniques can be used to detect chromosome
1081 territories, 3D fluorescent *in situ* hybridization (FISH) is most useful to study this level of organization, as
1082 it provides direct spatial information^{7,27,41}. Compartments and hubs are usually studied with Hi-C, GAM,
1083 split-pool recognition of interactions by tag extension (SPRITE) and several multiplexed super-resolution

1084 FISH techniques^{14,15,123,27,35,41,44-48}. At the more functional, 1 Mb scale, most informative are capture Hi-C
1085 (cHi-C), Capture-C, micrococcal nuclease chromosome conformation assay (Micro-C) and super-resolution
1086 FISH approaches^{21-26,91}. On the other hand, protein-driven enrichment techniques such as Hi-C chromatin
1087 immunoprecipitation (HiChIP), chromatin interaction analysis by paired-end tag sequencing (ChIA-PET),
1088 proximity ligation assisted chromatin immunoprecipitation (PLAC-seq) and chromatin-interaction analysis
1089 via droplet-based and barcode-linked sequencing (ChIA-Drop) can be used to study different levels of
1090 folding depending on whether the protein is associated with relatively local folding (e.g., MED12 in E-P
1091 contacts), more long-range contacts (e.g., Polycomb proteins) or inter-chromosomal hubs (e.g., LHX2,
1092 LDB1)^{31-33,102,192,215-218}. CNDs have been discovered only recently. Owing to their stochastic nature and to
1093 inter-cell variability, only super-resolution microscopy FISH has so far been able to detect CNDs^{43,57,117}.
1094 DNA adenine methyltransferase identification (DamID), Genomic loci positioning by sequencing (GPSeq),
1095 Tyramide signal amplification (TSA).

1096

1097 **Figure 2 Main C-based methods for interrogation of 3D genome organization.**

1098 For the application of every chromosome conformation capture (C)-based method, chromatin must first
1099 be cross-linked, either with one cross-linker (in most methods) or with two cross-linkers in the case of
1100 micrococcal nuclease C (Micro-C) and chromatin interaction analysis by paired-end tag sequencing (ChIA-
1101 PET)^{85,91,102}. Hi-C derivatives then use digestion by restriction enzymes (RE) to fragment the chromatin, fill-
1102 in fragment ends with biotin and perform ligation, all *in situ*. The sample is then sonicated, de-crosslinked
1103 and enriched for informative fragments through biotin pulldown; the resulting chromatin-fragments
1104 library is subjected to amplification and sequencing⁸⁵. Hi-C chromatin immunoprecipitation (HiChIP)
1105 differs from this standard protocol by introducing an immunoprecipitation (enrichment) step just after
1106 ligation, whereas capture Hi-C (cHi-C) includes an enrichment step involving hybridization to RNA baits
1107 that represent a genomic region of interest, followed by pulldown that is performed on the final Hi-C
1108 library^{26,32}. The resulting libraries will therefore be enriched either for all genomic contacts (in the case of
1109 Hi-C), for chromatin contacts at genomic regions where the protein of interest binds (HiChIP), or for
1110 contacts at a specific region of interest (cHi-C). Micro-C follows an almost identical procedure as Hi-C, with
1111 distinctions in the steps of fixation and of digestion, in which the restriction enzymes used in the Hi-C
1112 protocol are replaced with micrococcal nuclease (MNase)-mediated chromatin fragmentation⁹¹. This
1113 digestion modification results in a more uniform fragmentation of the genome, thereby allowing Micro-C
1114 to achieve higher resolution of local contacts. ChIA-PET provides similar information to HiChIP, but in ChIA-
1115 PET immunoprecipitation is performed immediately after fixation and sonication; and while the sample is

1116 still on the pull-down beads, a linker with biotin is added and ligation is performed on-beads in order to
1117 reduce the amount of random ligation products¹⁰². Concomitant with adapter addition, the sample is de-
1118 crosslinked and fragmented using a transposase (Tn5). Finally, the sample is enriched for informative
1119 fragments through biotin pulldown and undergoes library amplification and sequencing. DSG,
1120 disuccinimidyl glutarate; EGS, ethylene glycol bis(succinimidyl succinate).

1121

1122 **Figure 3 Main ligation-independent methods for interrogation of 3D genome organization.**

1123 Ligation-independent methods have been developed in order to study multi-way contacts that are
1124 inaccessible to ligation-based methods. The most common ligation-independent techniques are split-pool
1125 recognition of interactions by tag extension (SPRITE), chromatin-interaction analysis via droplet-based and
1126 barcode-linked sequencing (ChIA-Drop) and genome architecture mapping (GAM)^{35,41,192}. In all three,
1127 nuclei are fixed. Then, SPRITE and ChIA-Drop proceed with sonication and chromatin digestion by DNase
1128 I. From this point on, the two techniques take advantage of different approaches to retaining information
1129 on multi-way contacts. SPRITE utilizes a split-and-pool strategy, in which every sample is split, barcoded
1130 and pooled together five times. This results in unique barcoding of all fragments that crosslinked together
1131 and thus to the identification of DNA sequences that were involved in the same multi-way contacts³⁵.
1132 ChIA-Drop uses microfluidics to produce a droplet carrying a unique barcode, adapters and material for
1133 DNA amplification reactions used to label a single chromatin-interaction knot, thereby allowing the
1134 identification of all DNA sequences that have been crosslinked together. GAM utilizes a completely
1135 different strategy to assay genome architecture and is suitable for investigating multi-way contacts,
1136 higher-order chromatin structures as well as more local contacts. In GAM, the sample is fixed, embedded
1137 in sucrose and cryo-sectioned to obtain thin slices, from which individual nuclear slices are laser micro-
1138 dissected. Genomic DNA is then extracted from a single-nucleus slice followed by whole-genome
1139 amplification and sequencing. The data obtained from the different sections of a nucleus are pulled
1140 together and interactions are identified as DNA sequences that co-segregate more often than others.

1141

1142 **Figure 4 Microscopy and FISH-based methods for 3D genome investigation.**

1143 **a)** Super-resolution microscopy techniques overcome the diffraction limit using different approaches. In
1144 structured illumination microscopy (SIM), the sample is exposed to a series of non-uniform illumination
1145 from different angles and axial phases¹⁰⁷⁻¹⁰⁹. The resulting light pattern is analysed by Fourier
1146 transformation to achieve a final mathematical reconstruction of the image, which improves lateral and
1147 axial resolution by two-fold. Single molecule localization microscopy (SMLM) uses low excitation energy

1148 that causes a stochastic excitation of photoswitchable fluorophores and allows for the precise localization
1149 of the centre of emission^{112–115}. Sequential images of the sample are taken, in which fluorophores turn
1150 either bright or dark and the final image is created by a superposition of all imaging cycles. In most
1151 practical applications, this method yields a lateral resolution of up to 20 nm. Stimulated emission
1152 depletion (STED) uses stimulated emission depletion through the combination of two lasers: an excitation
1153 laser illuminates the sample in the middle and a doughnut-shaped depletion beam depletes the
1154 surrounding signal^{118,119}. In practice, STED reaches lateral resolution of about 50 nm and axial resolution
1155 of about 80 nm to 600 nm. **b)** Oligopaints are fluorescently labelled synthetic DNA oligonucleotides that
1156 can be combined with fluorescence *in situ* hybridization (FISH) to label, visualize and measure the
1157 distances between genomic regions¹²². **c)** Oligopaints can be further modified to allow sequential imaging
1158 by incorporating sample bleaching or automatized microfluidics with microscopy. Following every round
1159 of imaging, the oligonucleotides are washed out before proceeding with a new round of hybridization and
1160 imaging^{44–48,123}. Distant genomic regions are imaged simultaneously using different detection
1161 fluorophores while their neighbouring regions are imaged in the next round, thereby allowing chromatin
1162 tracing. **d)** In oligo fluorescence *in situ* sequencing (OligoFISSEQ), oligonucleotides contain barcodes that
1163 can be read through hybridization with a set of specific primers, to which fluorescently labelled
1164 dinucleotides are ligated. The fluorescent signal is then imaged and cleaved off. This process is repeated
1165 until the barcode is read in full (inferred from the specific combination of fluorophores)^{14,15,219} gDNA,
1166 genomic DNA; PSF, **point spread function**; SPDM, spectral position determination microscopy.

1167

1168 **Figure 5. Computational modelling of the 3D genome.**

1169 Two main computational strategies have been devised to investigate chromosome folding and the
1170 mechanisms driving it: a bottom-up strategy (polymer modelling) and a top-down strategy (restraint-
1171 based modelling)^{137,157,158}. **a)** Polymer modelling mimics the physical behaviour of the chromatin fibre that
1172 is represented as series of monomers (beads) on a string and in which attraction or repulsion forces can
1173 be attributed based on first principles. The behaviour of the polymer is influenced by non-specific
1174 restraints, e.g. imposing that the beads must be self-avoiding, by specific restraints, such as cohesin-
1175 dependent loop extrusion that is blocked at convergent CTCF sites, and by specific interactions, such as
1176 attraction between beads sharing the same epigenetic modifications. Such strategy can produce virtual
1177 contact maps that can be compared with Hi-C interaction matrices. This process is reiterated until the
1178 input parameters can reconstitute a simulation that optimally recapitulates experimental maps. Polymer
1179 modelling allows to infer the mechanisms and estimate the forces that are necessary or sufficient to

1180 achieve the chromatin conformation closest to the experimental data¹⁵⁹. Beads of different colours
1181 indicate regions carrying different epigenetic modifications **b)** Restraint-based modelling uses
1182 experimental data such as Hi-C contact maps or fluorescent in situ hybridization (FISH)-measured nuclear
1183 distances to infer all the spatial restraints necessary to determine the structure of the genomic region of
1184 interest. The resulting model is an optimally reconstructed 3D folded chromatin fibre that gives
1185 information on spatial positioning of chromatin regions.

1186
1187 **Figure 6. Live microscopy using CRISPR-dCas9 to study the 4D genome.**

1188 Live-cell microscopy is one of the few approaches that informs on the dynamics of chromatin contacts.
1189 Multiple different methods allow the study of chromatin in 4D (changes of 3D chromatin structure over
1190 time)^{50,204–207}. Here, we illustrate two such methods based on nuclease-dead Cas9 (dCas9) recruitment
1191 to the chromatin: chimeric array of gRNA oligonucleotides (CARGO) and Live FISH. **a)** CARGO utilizes a
1192 single plasmid encoding multiple guide RNA (gRNAs) targeting the same genomic region of interest, for
1193 example an enhancer, in order to amplify the fluorescence provided by dCas9-EGFP⁴⁹. The region of
1194 interest can be imaged over time to study its spatial dynamics. **b)** Live FISH utilizes two fluorescently
1195 labelled gRNAs (one red (Cy3) and one green (A488)) targeting different regions of interest²⁰⁶. The two
1196 regions can be labelled and imaged simultaneously, which allows tracking the dynamics of the regions
1197 over time. D, distance.

1198

1199 **Box 1. Overview of proximity-ligation-based methods for mapping chromatin interactions**

1200 The development of chromosome conformation capture (3C), which detects pair-wise interactions
1201 between select loci ('one-vs-one') through nuclear-proximity ligation in combination with semi-
1202 quantitative PCR, marks the onset of the eponymous C-based techniques in the early 2000s^{17,18,29}. Using
1203 3C, the locus control region of the β -globin locus was shown for the first time to form chromatin loops
1204 with and thus to activate its promoter, and to form an active chromatin hub that dynamically follows
1205 transcription during differentiation and is stabilized by transcription factors^{220–223}. However, 3C is low-
1206 throughput and cannot successfully detect long-distance contacts. Circular 3C (4C) overcame these
1207 limitations by using primers in order to detect genome-wide contacts formed by a single 'viewpoint' to
1208 (one-vs-all)^{19,224,225}. Later, 4C was combined with next generation sequencing, and was used to describe
1209 the dynamics of chromatin contacts during development^{226–229}. However, the most influential technique
1210 in 3D genome organization research is Hi-C, in which the DNA interactome of the entire genome is assayed
1211 (all-vs-all)²⁷. Hi-C led to the identification of genomic compartments and topologically associating

1212 domains and to the development of the **loop-extrusion model** ^{27,53,130,54,55,58,59,61,64,67,85}. Finally, several
1213 techniques were developed, which combined Hi-C with chromatin immunoprecipitation, thereby allowing
1214 the interrogation of chromatin contact frequencies based on the presence of a specific protein ^{31–33}. It is
1215 important to note that in all C-based techniques that include protein enrichment, the mapped contacts
1216 are probabilistic rather than deterministic features and it is impossible to predict how and whether these
1217 contacts will translate into function.

1218 Together, proximity-ligation based techniques fueled most of the discoveries in 3D genome organization
1219 research during the past fifteen years. The ease of application of these techniques is anti-correlated with
1220 the richness of data obtained from them, which is likely the reason why so many different C-based
1221 technique adaptations exist. With the recent development of microscopy-related techniques, the
1222 proximity ligation-based data is even more valuable as it will offer an imaging-complementary information
1223 that is invaluable for achieving a better understanding of genome folding.

1224

1225 **Box 2 Manipulation of genome architecture**

1226 A fundamental question that has been very difficult to address in the past is whether genome architecture
1227 changes that are detected in different cellular conditions are a cause or a consequence of changes in gene
1228 expression ^{65,230}. Use of Hi-C in *Drosophila melanogaster* lines that carry a set of known genomic
1229 alterations has allowed to address this question. Intriguingly, genome topology changes stemming from
1230 chromosomal inversions or other mutations can be buffered to a large extent, such that gene expression
1231 remains robust in most, although not all genes²³¹. Whereas this analysis used pre-existing mutant lines,
1232 CRISPR–Cas technology has also been used extensively to manipulate genome architecture. Inducing
1233 specific mutations at critical genome architecture regulatory regions has shown that genome organization
1234 into topologically associating domains with specific boundaries contributes to correct cell-type and tissue-
1235 specific gene regulation^{213,232}. In *D. melanogaster*, the deletion of specific chromatin-loop anchor
1236 sequences or the insertion of boundary elements that prevent loop formation, showed that **Polycomb-**
1237 dependent genomic loops can contribute to gene silencing during development ²³³. The same approach
1238 showed that CTCF-binding sites are required for correct insulation of gene expression and that their
1239 deletion activates an oncogenic gene-expression program ²³⁴.

1240 CRISPR–Cas was used not only to edit the genome, but also the epigenome, as in the case of using a fusion
1241 of nuclease-dead Cas9 (dCas9) with DNA (cytosine-5)-methyltransferase 3A, which targets DNA
1242 methylation to specific regions and displaces CTCF binding²³⁵. Another powerful experimental tool
1243 combines CRISPR–Cas with optogenetics to induce chromatin looping upon stimulation with blue light and

1244 study its functional consequence²³⁶. Finally, the CRISPR-genome organization (CRISPR-GO) system enables
1245 inducible and reversible repositioning of dCas9-targeted genomic regions in the nucleus; CRISPR-GO was
1246 shown to reposition a locus of interest to the nuclear lamina, thereby perturbing its function, and to **Cajal**
1247 **bodies [G]** and **PML bodies**²³⁷. This system could be used to target loci to other nuclear compartments in
1248 order to study the functional consequences of their relocation. Combined with the many experimental
1249 tools that enable measuring the effects of perturbation of genome structure and function, these
1250 techniques will be crucial for differentiating between cause and consequence of 3D genome organization
1251 and gene regulation.

1252

1253

1254 **GLOSSARY**

1255 **First principles**

1256 Basic building blocks of knowledge that cannot be deduced from any other preposition used for
1257 mathematical modeling of polymer behaviors.

1258 **Airy diffraction pattern**

1259 A diffused circle surrounded by rings of decreasing intensity generated when a laser passes through a
1260 circular opening.

1261 **Cajal bodies**

1262 Nuclear bodies of 0.3 to 1 μ m in size, containing RNAs and proteins and involved in RNA-related metabolic
1263 processes.

1264 **Dendrimer crosslinking**

1265 A procedure in which formaldehyde crosslinking can be followed or replaced by crosslinking with
1266 dendrimers, i.e. highly ordered, branched polymeric molecules of different sizes.

1267 **Diffraction limit**

1268 The points where two Airy patterns are too close to be distinguishable.

1269 **Lamina associated domains**

1270 Chromosome domains associated to the nuclear lamina in the 3D nuclear space.

1271 **Loop extrusion model**

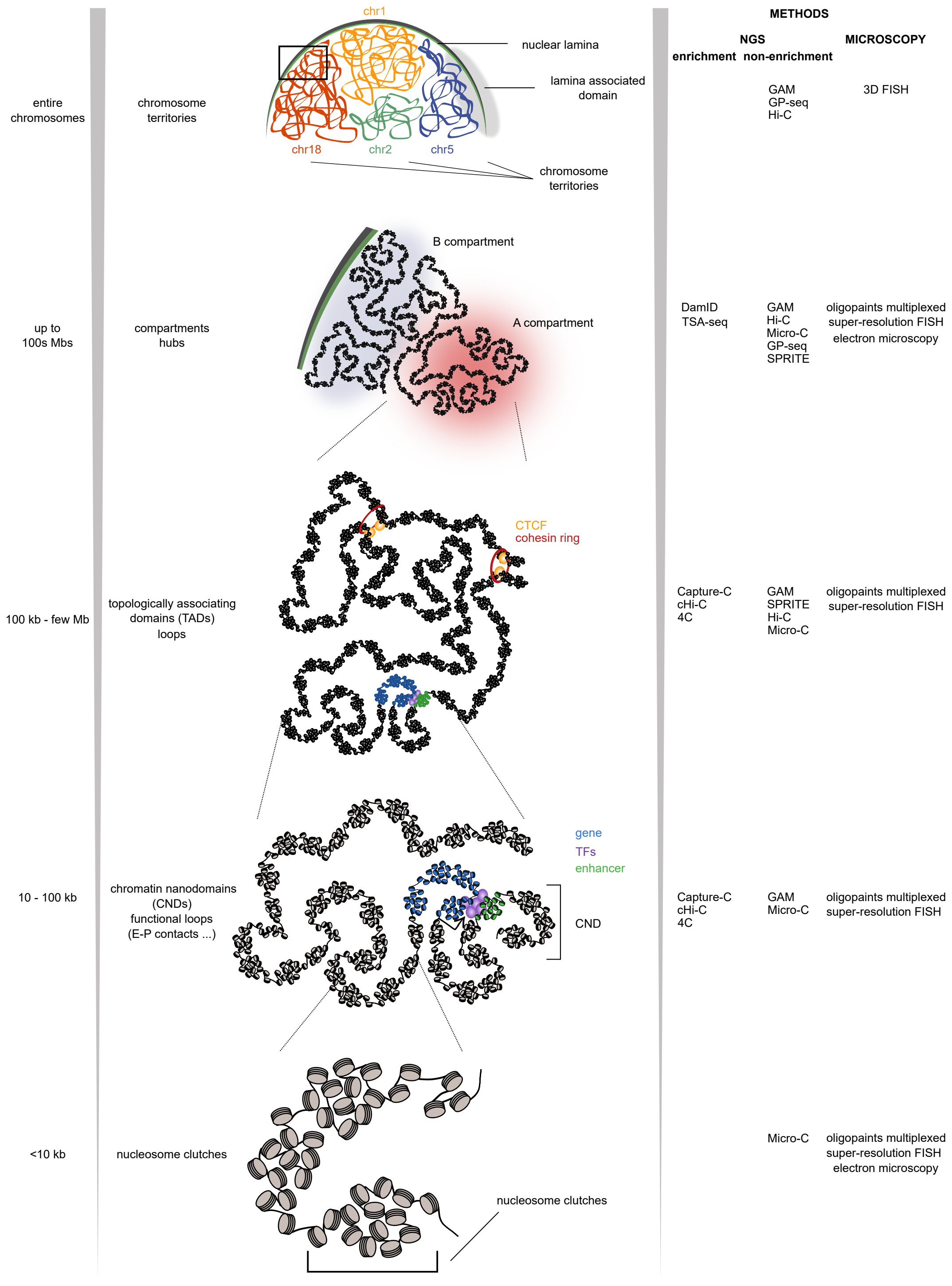
1272 Model suggesting that motors such as cohesin or condensin form a ring around chromatin and use the
1273 energy of ATP to slide through it while extruding the intervening region.

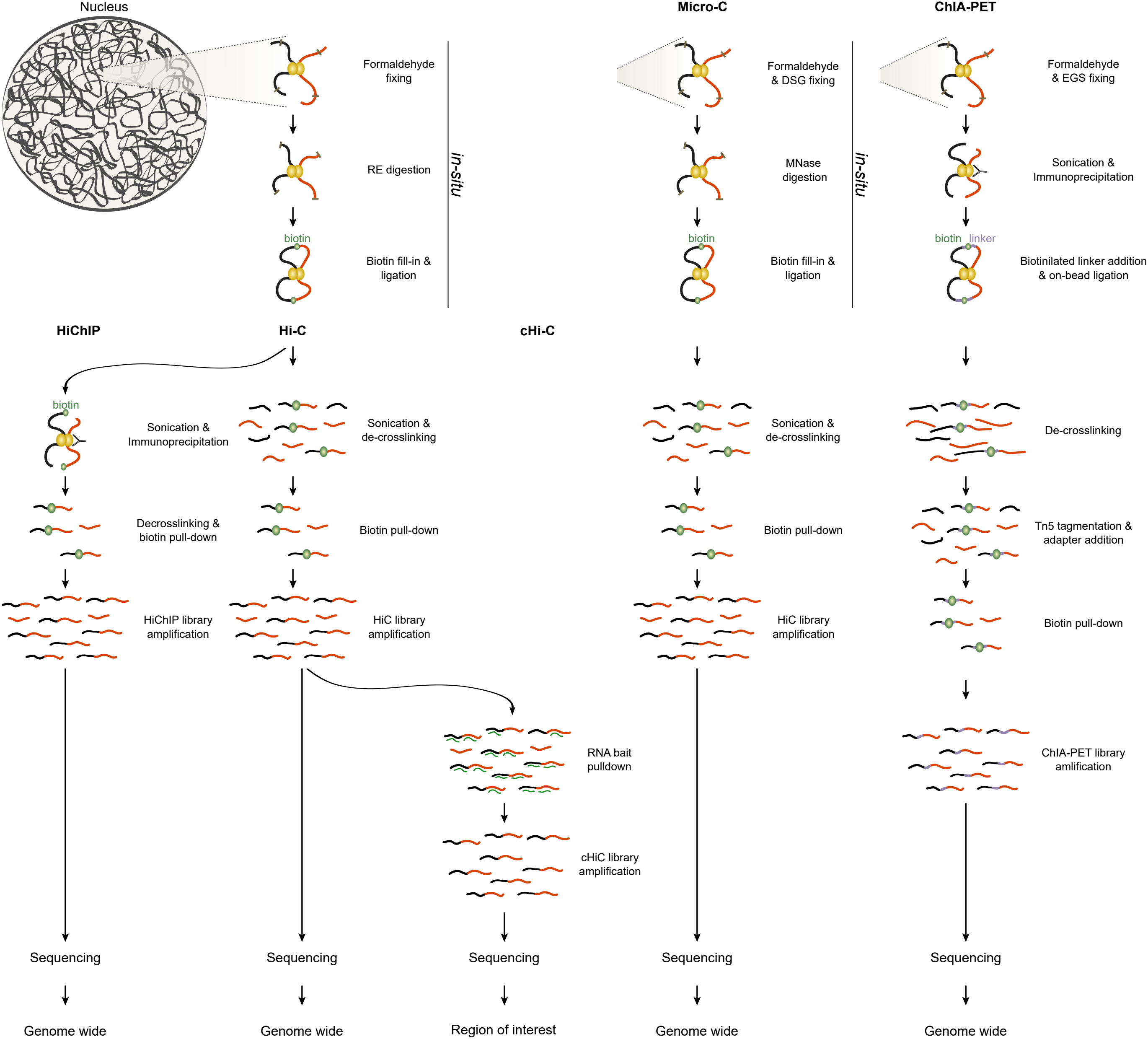
1274 **Multi-way contacts**

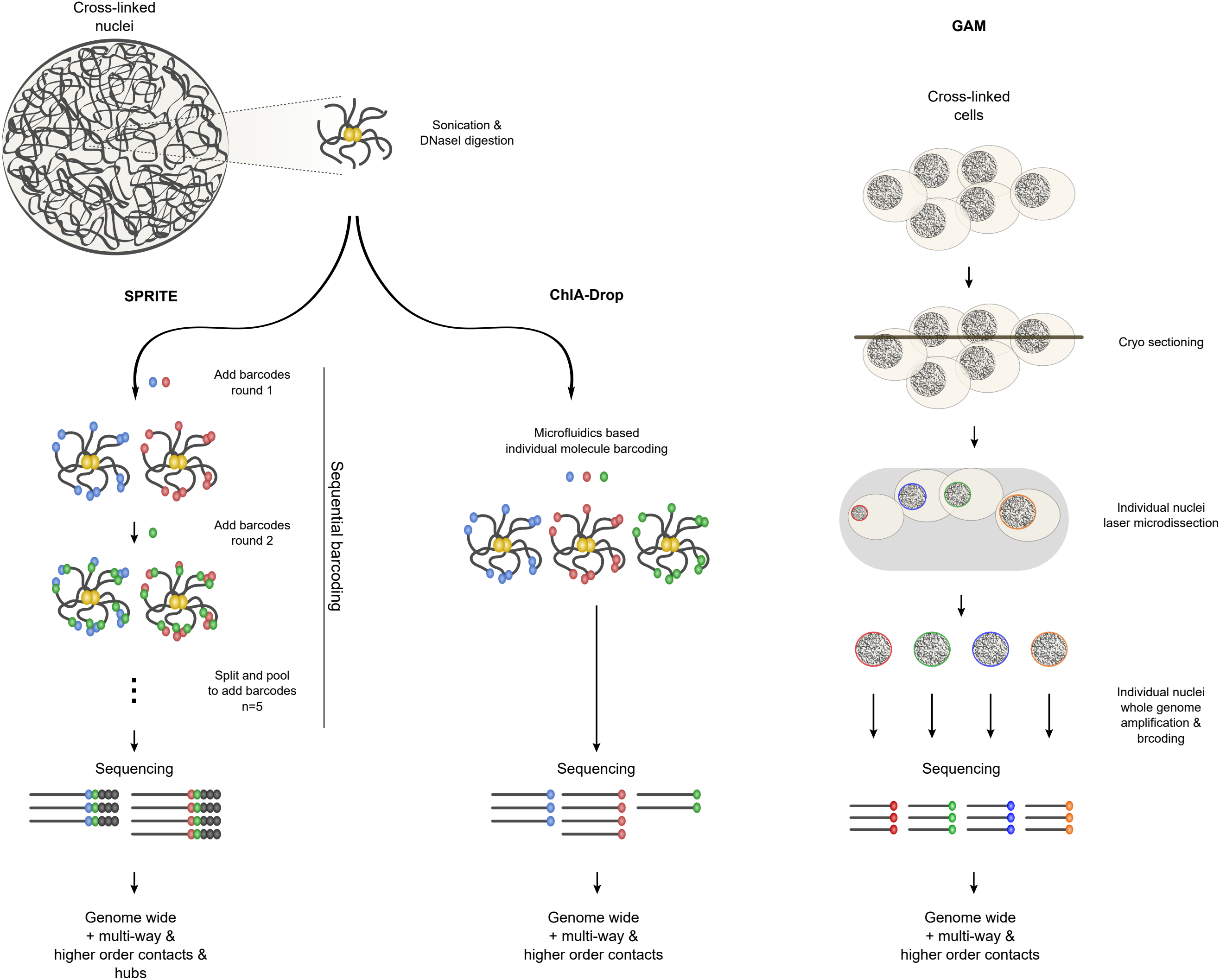
1275 Chromatin contacts involving more than two chromatin fragments.

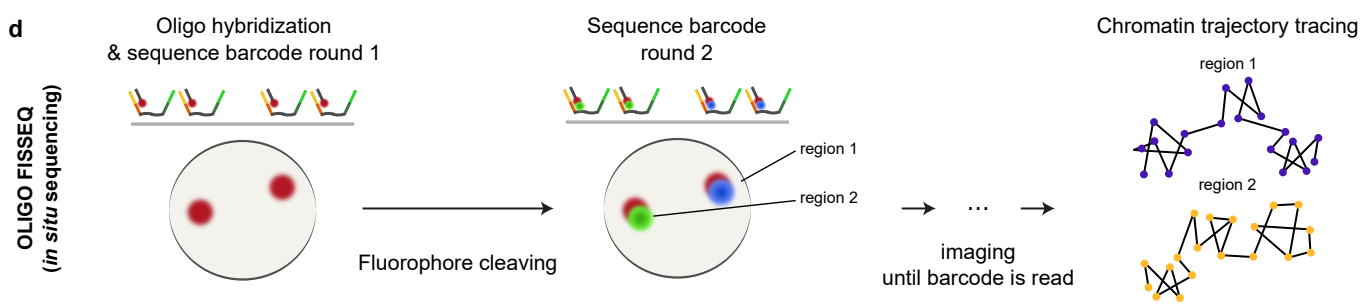
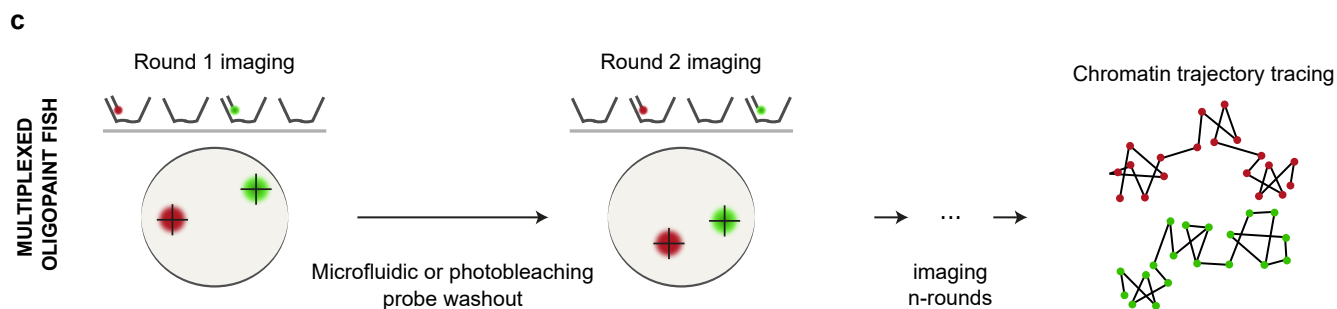
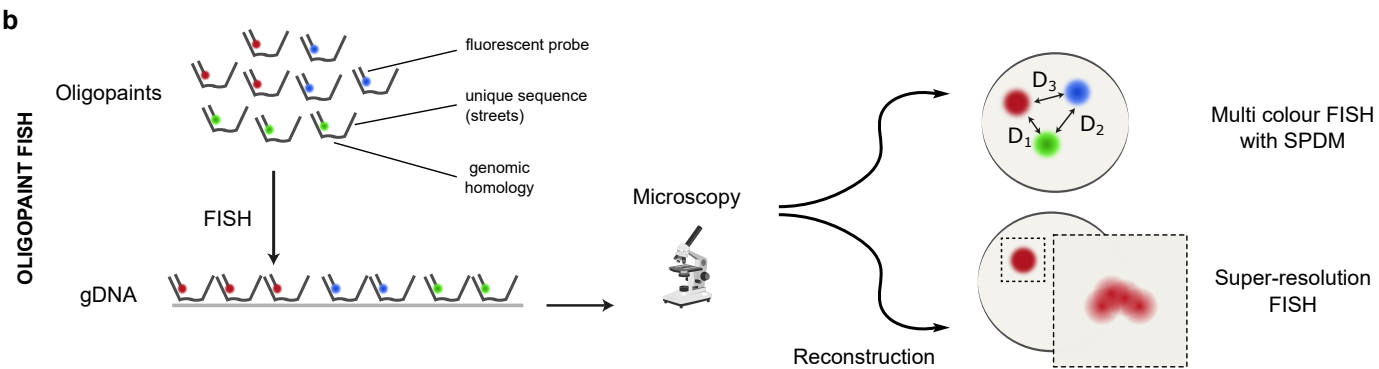
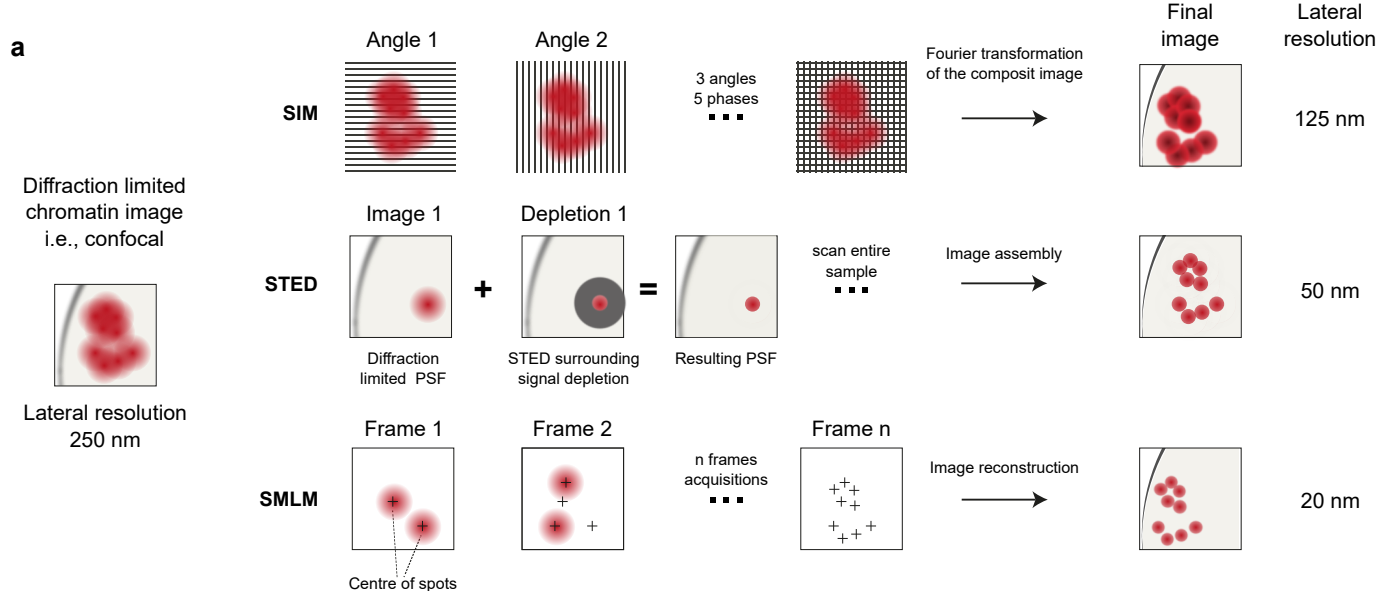
1276 **Nuclear speckles**

- 1277 Nuclear foci enriched in pre-mRNA splicing factors located in the nucleoplasm of eukaryotic cells.
- 1278 **Point spread function (PSF)**
- 1279 The response of an imaging system to a point object. If the object is below the microscope resolution it
1280 will appear larger than it really is..
- 1281 **PML bodies**
- 1282 Nuclear bodies of 0.1 to 1 μ m in size, containing many components, including the promyelocytic leukemia
1283 protein (PML) and frequently associated to Cajal bodies.
- 1284 **Polycomb**
- 1285 An evolutionarily conserved group of proteins involved in the regulation of a large group of target genes.
- 1286 **Sub-diffractive point spread function**
- 1287 A point spread function of smaller size than that generated by diffraction-limited systems.
- 1288 **Tyramide signal amplification**
- 1289 A method enabling sensitive detection of low-abundance molecules in fluorescent immunocytochemistry
1290 applications.



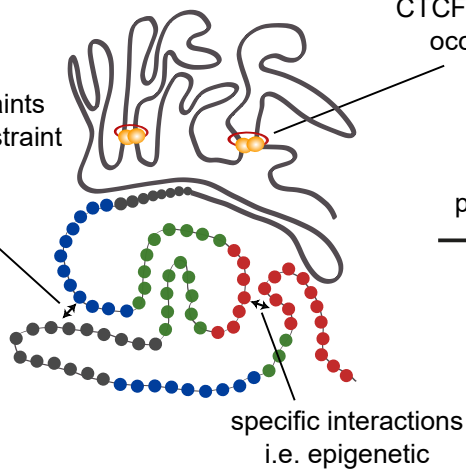
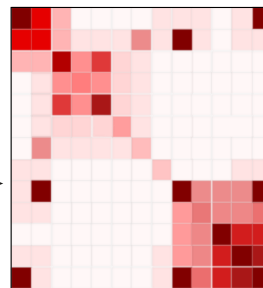
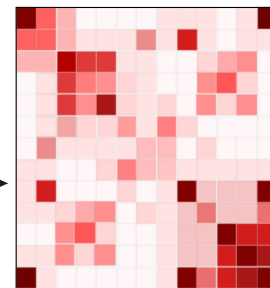






a

POLYMER MODELING

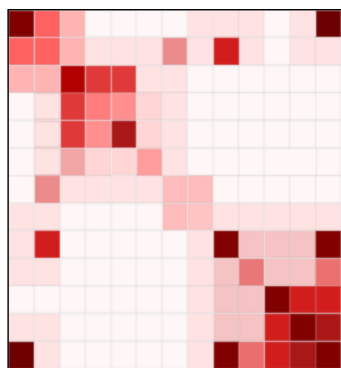
non-specific restraints
i.e. topological constraintsimulation of
polymer foldingcomparison
with experiment

repeat to find the conditions that best recapitulate the experiment

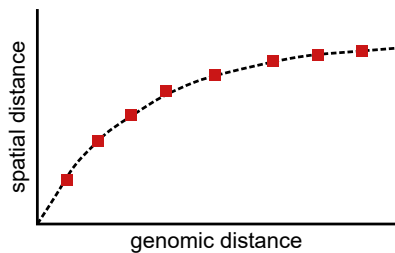
b

RESTRAINT MODELING

HiC data



FISH data

input experimental
constraints
i.e. contacts, distances

conformational ensemble

