The spatial organization of transcriptional control

Antonina Hafner and Alistair Boettiger ✉

Abstract | In animals, the sequences for controlling gene expression do not concentrate just at the transcription start site of genes, but are frequently thousands to millions of base pairs distal to it. The interaction of these sequences with one another and their transcription start sites is regulated by factors that shape the three-dimensional (3D) organization of the genome within the nucleus. Over the past decade, indirect tools exploiting high-throughput DNA sequencing have helped to map this 3D organization, have identified multiple key regulators of its structure and, in the process, have substantially reshaped our view of how 3D genome architecture regulates transcription. Now, new tools for high-throughput super-resolution imaging of chromatin have directly visualized the 3D chromatin organization, settling some debates left unresolved by earlier indirect methods, challenging some earlier models of regulatory specificity and creating hypotheses about the role of chromatin structure in transcriptional regulation.

Topologically associating domains
(TADs). These are defined on population-level contact-frequency maps as domains of higher interaction frequency within a region than between regions.

Department of Developmental Biology, Stanford University, Stanford, CA, USA.

✉*e-mail: boettiger@stanford.edu*

https://doi.org/10.1038/s41576-022-00526-0

In animal genomes, which genes are turned on or off is often controlled by distal *cis*-regulatory elements (CREs), particularly enhancers (BOX 1). These elements can activate or silence transcription of genes and yet are frequently positioned thousands to millions of base pairs from the transcription start sites (TSSs) of the gene(s) they regulate[1–3]. It is estimated that the average mammalian gene in any given cell type is under the influence of four distinct enhancers at a time[4], although this distribution is broad, with some developmentally important genes controlled by tens of enhancers[5,6] and other genes largely under the control of promoter proximal elements. The specificity exhibited by these enhancers is non-autonomous: if moved to a new locus they will generally activate any gene in their vicinity. Such 'enhancer hijacking' events have been increasingly identified as a driver of pathogenic behaviour of structural variants[7,8], as well as an essential enabling feature of numerous genomic assays including enhancer traps and massively parallel reporter assays such as STARseq[8,9]. Thus, determining how communication between these distant genomic regions is achieved and regulated is essential to any detailed understanding of transcriptional control in healthy development and its dysregulation in many pathologies.

In recent years, considerable progress has been made toward understanding the properties and mechanisms of this transcriptional control at long range[10–14], stimulated by an explosion of new technologies[10,13] and concerted investments such as the National Institute of Health's 4D Nucleome project[15]. In this Review, we explain these advances and discuss their relevance to our understanding of transcriptional control, including new models of transcriptional regulation, featuring enhancer–promoter (E–P) loops, topologically associating domains (TADs), and pre-formed and dynamic architectures. As these advances have been extensively reviewed elsewhere[10–14,16–19], we will concentrate on discussing those results that have led to conflicting interpretations owing to intrinsic limitations of the methodologies. Then, we introduce the complementary strengths and limitations of emerging microscopy-based methods for mapping chromatin nanostructure and studying *cis*-regulation in single cells. We discuss how they have contributed to a revised understanding of the nature of TADs and E–P contacts, and the mechanisms by which these features influence transcriptional regulation. In the interest of brevity, and to contrast with sequencing approaches, we focus on fixed-cell, super-resolution mapping approaches rather than live cell imaging. We close by highlighting some of the outstanding questions in gene regulation.

Indirect measures of 3D organization

Early evidence of a role for chromatin folding in transcriptional control. It has long been believed that distal enhancers must achieve some degree of spatial proximity with the (core) promoter they control, both to allow proteins bound to the enhancer to interact with those bound (or trying to bind) to the promoter and to explain why the regulation happens only in *cis* and not in *trans*. It has remained more controversial whether this proximity is a regulated step in transcriptional control. If the mobility of chromatin is high, contact frequency

Box 1 | **Terminology for transcriptional regulatory regions**

The 'enhancer' was defined as a sequence element capable of stimulating transcription in a distance-independent and orientation-independent manner[20]. The best-studied examples of enhancers contain specific combinations of binding sites for both transcriptional activator proteins and transcriptional repressor proteins, both of which contribute to the cell-type specificity of expression[1,8,163]. The broader concept of 'distal *cis*-regulatory element' (CRE) has been offered as a more accurate replacement of the word enhancer in discussing long-range transcriptional control generally, to emphasize that these sequences are not necessarily exclusively involved in gene activation or merely augmenting or enhancing some other more primary mode of regulation. However, as the activating effects of distal elements are much better studied than the repressive roles, we will retain here the tradition of referring primarily to 'enhancers', while acknowledging that repressive regulation may require similar logic for distal control.

The 'promoter' is understood to mean the regulatory sequence upstream of the coding sequence of a gene that controls when and where it is transcribed. In simple single-celled organisms, such as *Escherichia coli*, the complete regulatory sequence of a gene is almost always upstream of the coding sequence and within several kilobases of the transcription start site (TSS). In animal genomes, including human genomes, the region approximately 1–5 kilobases (kb) upstream of the TSS also frequently contains binding sites for cell-type-specific transcription factors to induce transcription in the appropriate cell type and is also called a promoter. However, many genes are heavily regulated by multiple, more distal elements, which may be upstream, downstream or intronic, and may lie thousands to hundreds of thousands of base pairs away, including on the far side of neighbouring genes or even in the introns of unrelated genes[2,4,8]. Many of the distal sequences of the full promoter (that is, regulatory regions lying 1–5 kb upstream of the TSS), are equally capable of activating transcription from downstream positions and in an orientation-independent manner — thus meeting all the criteria of enhancers and CREs and in contrast to the behaviours observed for bacterial promoters. In light of this, it is better to distinguish these separate functions as a TSS-proximal enhancer and a 'core promoter'.

The core promoter, typically less than 150 base pairs in length, is defined as necessary but not sufficient to drive gene expression in a chromatin (in vivo) context. This sequence always includes the TSS, and is typically enriched for one of several short motifs associated with high transcription potential, such as the initiator, downstream promoter element (DPE), or TATA box, although some degree of core promoter activity may be achieved without a consensus match to any of these elements[1,164,165]. The core promoter lacks binding sites for the cell-type-specific transcription factors needed to drive gene expression in a chromatin context. These factors may bind immediately upstream of the core promoter (part of the 'full length' promoter), or they may bind thousands or hundreds of thousands of base pairs distal to it.

The functional element of these definitions for enhancer and core promoter (for example, necessary but not sufficient for transcription) makes them especially useful, but also partly ambiguous (for an excellent discussion see REF.[166]). Far from the simplistic on/off view, whether or not transcription occurs is dependent on the sensitivity of the assay and an element that seems to be insufficient for transcription in one assay may have some basal expression rate in another. Although it is conceptually useful to define enhancers as the sites of cell-type-specific regulatory control and core promoters as the sites of transcription initiation regulated by enhancers, it should be noted that this is an over-simplification. Binding sites for cell-type-specific regulation may indeed be found mixed up among other core promoter sequences[1]. A promoter of one gene may act as an enhancer for another[167].

may not be rate limiting to protein–protein interactions between an enhancer and promoter. But if chromatin diffusion is slow, then it is possible that mechanisms exist to accelerate the frequency and selectivity of E–P interactions. Genetic experiments provide increasing evidence for such regulated spatial proximity. For example, from their first discovery, enhancers from one gene were shown to activate expression from multiple different core promoters, as long as the enhancer was placed on the same plasmid or within about ten kilobases of the promoter[20–23] (note that the genetic tools available 20 years before the advent of Cas9-based gene editing largely precluded analysis of larger distances). These experiments suggested that contact frequency may be key to regulatory specificity at long range, rather than biochemical compatibility codes between the enhancer and promoter serving as the primary determinant[20–23]. The discovery of insulator elements, sequences which interfered with this E–P communication, further suggested that contact frequency may be regulated in order to provide specificity[24–26]. However, the first critical advance in directly testing the role of regulated 3D organization in guiding transcription was the ability to map the spatial proximity of all genome sequences relative to one another. This has been chiefly achieved with chromatin conformation capture (3C) assays, which cut the genome into fragments, ligate together the fragments that are in close physical proximity, and identify these ligation junctions via sequencing. Numerous reviews have already highlighted the immense technical progress in throughput of evolving 3C assays over the years (some of our favourites include REFS.[10,13,27–30]). Less has been written about the distinct aspects of these techniques that contribute to some contrasting views of 3D organization of transcriptional regulation, which we discuss below.

3C supports an E–P bridge model of chromatin organization. Early 3C approaches used PCR with primer pairs positioned at particular candidate loci to test whether any spatially mediated ligation events had occurred between the probes tested[31] (FIG. 1a). Many E–P pairs tested with this approach showed signal only in cell types in which the enhancer was bound by activating transcription factors (TFs) and the promoter was actively transcribed, and not in other cell types[32–37] (FIG. 1b). These observations led to a model of enhancer-mediated transcriptional regulation in which E–P proximity does not occur until the enhancer is active (that is, bound by activator TFs and/or has gained accessibility and histone marks associated with active enhancers)[38]. Intervening DNA is hopped over or looped out by the specific interaction of the enhancer- and promoter-bound factors (FIG. 1c). Anchored or positioned by cell-specific TFs, general TFs such as the Mediator complex and cohesin complex have been suggested to bridge the intervening sequence between enhancer and promoter that did not exhibit 3C contacts[34,39].

Despite supporting a compelling model of regulated E–P looping, there is a potential challenge with the interpretation of earlier 3C data. Detecting a ligated product between two DNA regions requires not only their spatial proximity, but also that both sites were successfully

Mediator complex
A multisubunit protein complex that is a general regulator of transcription by RNA polymerase II.

Cohesin complex
A multisubunit protein complex that mediates sister-chromatid cohesion in mitosis and is essential for topologically associating domain (TAD) formation.

cut in the first place. Active, TF-bound enhancers and actively transcribed promoters are much more accessible to restriction digestion. Thus, differences in accessibility only may, in theory, be sometimes mistaken for differences in spatial organization. Whereas biases in restriction site choice can be controlled by the use of different restriction enzymes and biases in amplification controlled by the use of multiple primers, biases in the ability of the sequence to be cut cannot readily be adjusted for in this approach. This limitation is shared to varying degrees by several elaborations of the procedure, such as 4C and promoter-capture Hi-C (for reviews see REFS.[10,13,27–30]). Convolving accessibility and proximity into a single measurement can be advantageous because it increases coverage of the experiment for enhancers and promoters, but only if the interpretation recognizes the convolution. Some Hi-C variations, such as Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) and HiChIP, also called Proximity ligation-assisted ChIP-Seq (Plac-seq)[40–44], intentionally enhance this bias and enrich for active chromatin regions; for example by including a selection step for histone H3 acetylated at lysine 27 (H3K27ac), a histone modification associated with active enhancers and promoters. These enrichment approaches have proved useful in predicting functional E–P links[41,42].

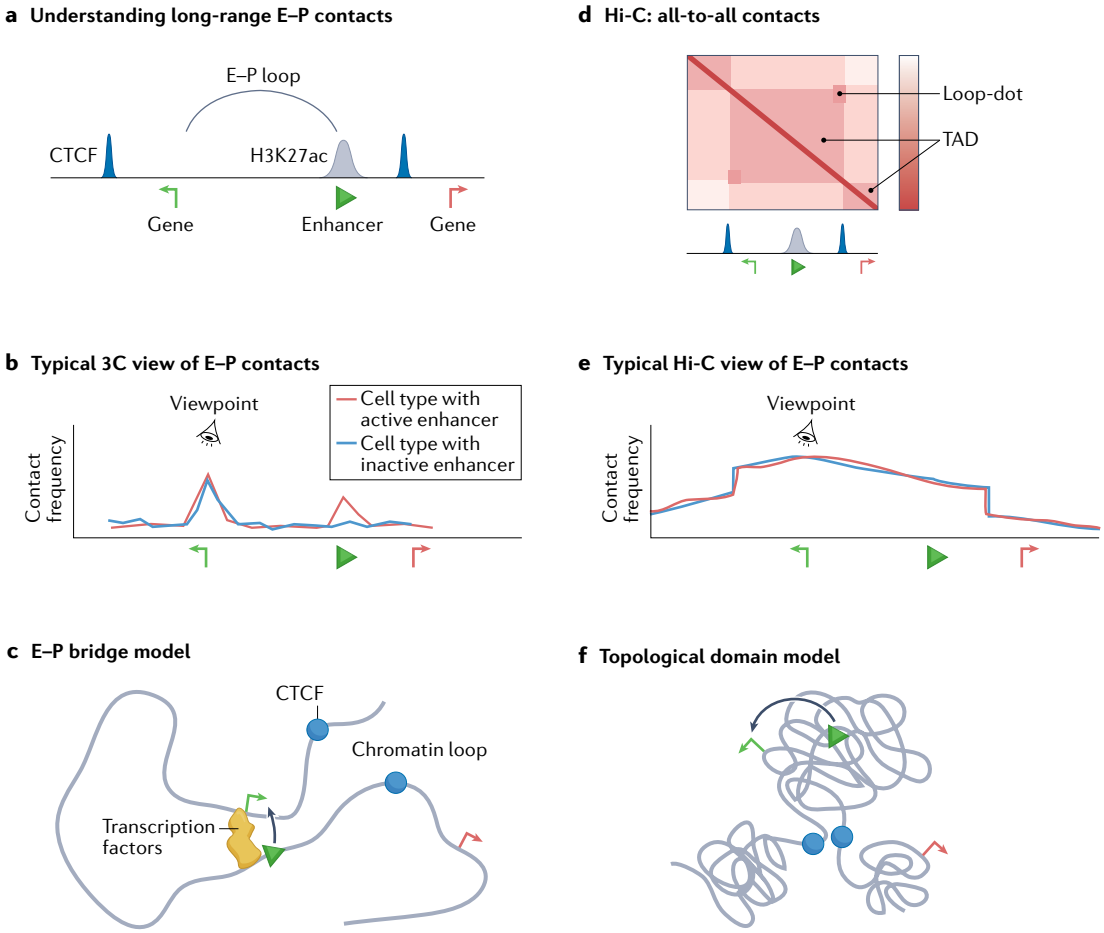


Fig. 1 | **Contrasting views of enhancer-mediated transcription regulation based on 3C and Hi-C data. a** | Schematic of a hypothetical genomic segment, containing two genes and one enhancer. The enhancer contains histone modifications associated with active enhancers (H3K27ac, histone H3 acetylated at lysine 27) and drives expression of the gene to the left, but not the gene to the right. Two peaks of CTCF binding are shown. **b** | The 3C method maps the frequency with which select genomic sites are in 3D proximity to a particular site of interest called the 'viewpoint', often a gene promoter of interest as shown. In 3C data, promoters and enhancers almost always appear as local peaks of contact frequency, elevated above intervening non-regulatory sequence. These local peaks are lost in a cell type in which the enhancer is no longer active. **c** | Depiction of the enhancer–promoter (E–P) loop view of regulation based on 3C data, in which transcription factors create a chromatin loop to anchor the enhancer to the promoter. **d** | Depiction of Hi-C data, in which boxes (topologically associating domains, TADs) are the primary feature, with occasional dots at TAD borders, called loops in much of the literature, but here referred to as 'loop-dots' to distinguish them from the polymer (chromatin) 'loops' shown in panel **c**. **e** | Hi-C contact frequency of all sequence elements with the promoter element, viewed as a line plot to compare to the 3C view shown in panel **b**; unlike 3C data, contacts can be seen in the regions between the enhancer and promoter. **f** | A schematic depicting the domain model of enhancer promoter specificity, popularized after the discovery of TADs from HiC data; enhancers and promoters that share a TAD may interact within a chromatin globule, but are physically separated from neighbouring TADs that sort into separate globules.

Hi-C supports a topological domain model of chromatin organization. The development of a non-selective, all-to-all contact approach called Hi-C[45] offered a new way of normalizing the accessibility bias and sequence-specific amplification bias inherent in earlier 3C methods. Given that all sequences should be equally represented in their proximity to something else because they are all present in equal abundance, and exploiting the fact that Hi-C data measures all-to-all interactions, the idea of matrix balancing by iterative normalization was introduced[46,47]. With this approach, although ligations between hypersensitive regions are still recovered at higher frequency than intervening regions, this computational correction adjusts these read counts downwards until they are as abundant as non-hypersensitive regions. As a result, Hi-C experiments provided a somewhat different view of the 3D landscape around E–P interactions (FIG. 1d,e). In contrast to the sparse peaks of specific contacts centred on promoters and enhancers generated by 3C (FIG. 1b), Hi-C analysis of all-to-all contacts revealed large domains with shared similar contact frequencies that change slowly as a function of genomic distance up to the edge of the domain. The edge of the domain is defined by a more abrupt drop of contact frequency, giving the domain a 'box-like' appearance in the contact map[48–50]. These domains were called TADs[48,50], and range in size from a few tens to a few thousands of kilobases. We note that although the changes in contact frequency appear as sharp transitions at the TAD border, the drop in contact frequency is typically only one- to two-fold (FIG. 1e).

Known regulatory domains (that is, regions spanning a promoter and all its known enhancers) were observed to frequently lie within a TAD and rarely to be separated by TAD borders. Furthermore, genes within the same TAD tended to have a correlated response to signals[48,51], whereas distinct spatial-temporal expression patterns could be observed for neighbouring genes separated by a TAD border[52,53]. Particularly instructive were genetic manipulations that deleted (or repositioned) these border sequences, which in some cases led to both a merger of the TADs and ectopic activation of the genes in one TAD by enhancers that, before fusion, were in the other TAD[7,18,53,54]. Unlike the view emerging from 3C studies, in which enhancers or genes themselves shape the 3D interaction landscape, TAD borders were more enriched for CTCF and cohesin binding than for binding of cell-type-specific TFs or for histone marks common to active enhancers[1,8]. Moreover, in contrast to what is seen in 3C data, the correspondence between a known regulatory domain and a TADs is observed even in cell types in which the gene and its enhancers are silent, with only minor differences in 3D organization compared with cell types in which they are all active (FIG. 1e). Prominent examples of TADs aligned to well characterized regulatory domains include the regulatory landscape of the *Shh* domain[55–58], the *EphA4* domain[59,60] and the *Sox9* domain[61]. In each case, the TAD contains all known enhancers up to a megabase from the TSS and can be detected in diverse cell types, including cells in which the enhancers and their target genes are silent.

Architectural loops sometimes overlap enhancers and promoters in Hi-C data. By and large, the structural view from Hi-C shows few dips in contact frequency within the megabase scale — that is, sequences that are more proximal to the promoter contact the promoter more frequently than (or at least as frequently as) more distal sequences do. However, there are notable exceptions. A subset of TAD border positions exhibit increased contact frequency relative to more proximal intra-TAD positions (FIG. 1d). These 'dots' are widely referred to as 'loops' in the Hi-C literature[47,62]; here, we will refer to them as 'loop-dots', to distinguish these features defined by a two-dimensional (2D) heatmap from a loop in a polymer (FIG. 1d). A subset of these loop-dots corresponds to well-studied E–P pairs, such as within the *Sox2* (REF.[63]), *Shh*[55,56] and *Myc*[47] regulatory domains. However, the specificity of loop-dots seems to be more dependent on CTCF than on enhancer activity. Cell types that retain CTCF binding, but in which the corresponding enhancer and promoter are inactive, frequently retain the loop, whereas experiments that delete CTCF binding sites in enhancers weaken or abrogate these loops[55,64].

It is not that Hi-C data suggest no role for E–P activity-directed interactions. A number of CTCF-marked loop-dots are cell-type-specific and only form concurrently with activity of associated genes and enhancers, particularly when CTCF sites are methylated in the inactive state; see, for example, REFS.[47,65]. A subset of enhancers and promoters that lack CTCF does exhibit loop-dots, for example[66], though they are generally weaker than CTCF-marked loop-dots. Furthermore, some constitutive loop-dots that overlap enhancers bound by CTCF become stronger when the enhancer and promoter are active, as observed at some loci involved in macrophage differentiation[66] and neural differentiation[67]. Nonetheless, most known enhancers are not loop-dot connected to their promoters in Hi-C data, in contrast to the general picture derived from 3C studies. If E–P interactions are more frequent than interactions with the intervening chromatin, the difference is generally slight compared with the contact differences that demarcate TADs and CTCF-marked loop-dots. These different scenarios suggest distinct mechanisms of E–P communication: the 3C peak-like view of contact suggests specific E–P tethers, in which the intervening sequence is much further away in 3D space (FIG. 1c), whereas the TAD view suggests that a promoter spends as much time scanning through non-enhancer sequences as it spends near functional enhancers within the TAD (FIG. 1f).

We note that the latest variant of the 3C family, called micro-C, has toned back this contrast. Micro-C uses micrococcal nuclease (MNase) rather than restriction enzymes to cut the chromatin prior to ligation, which allows for finer-scale distances to be probed, but with a noted preference for cutting accessible chromatin[68]. One micro-C study found mild contact bias among enhancers and promoters compared to non-regulatory intervening regions, and concluded these to be weak pause points for cohesin[69]. However, other micro-C[70,71] and targeted micro-C[72,73] approaches concluded that these E–P interactions were cohesin-independent, activity-associated contacts, supporting the original view from 3C data.

CTCF
The CCCTC binding factor (CTCF) is a zinc-finger transcription factor that is enriched at the boundaries of TADs.

Tethers
These are *cis*-regulatory elements that function to bring together distal DNA elements.

Phase separation
The process by which a mixture (such as oil and water) spontaneously separates into distinct phases, with molecules of one type separating from molecules of a second type.

Enthalpic
Interactions driven by the binding energy between molecules, such as homotypic interactions among chromatin states.

Loop extrusion
A model of how CTCF and cohesin are thought to form topologically associating domains (TADs), whereby cohesin is loaded onto the DNA and extrudes a loop until it is blocked by CTCF bound at the base of the loop.

Chromosome territories
Specific, largely non-overlapping areas in the nucleus that each chromosome occupies.

Notably, recent micro-C from *Drosophila melanogaster*, at particularly high resolution owing to the smaller genome, found that promoters and enhancers generally lack preferential contact, although a subset of E–P pairs reside near tethers that form more structurally stable contacts[74,75].

To determine to what extent enhancers and promoters are themselves major regulators of 3D structure (as suggested by 3C and some micro-C data) or instead occasional moderators of folds directed by other factors (as suggested by Hi-C and other micro-C data) it will help to turn to more diverse approaches. Orthogonal ligation-independent sequencing-based methods for mapping 3D interaction patterns have validated the existence of TADs but showed little evidence for E–P-directed folding. These approaches include split-pool recognition of interactions by tag extension (SPRITE)[76], which uses a split-pool barcoding approach instead of ligation to identify regions close in 3D space, and genome architecture mapping (GAM)[77,78], which bypasses both restriction digestion and ligation by implementing whole-genome amplification from ultrathin nuclear sections (for reviews see REFS.[10,13]). Both methods currently have lower resolution than 3C, Hi-C or micro-C, so interactions between more proximal E–P pairs or weak longer-range E–P loop-dots would be missed. With this caveat in mind, Hi-C and its ligation-free alternatives collectively suggest that the domain-like structural signature is stronger than the contribution of E–P interactions to 3D organization.

Sequencing-based studies raise new questions. As for all major advances, the new perspectives on transcriptional regulation arising from 3C, Hi-C and related methods opened as many new questions as they answered. Of particular interest is how residing in the same TAD contributes to E–P specificity. Many viewed and depicted TADs as spatially segregated globules, with sequence elements generally equidistant from other sequences within the same TAD and well separated from neighbouring TADs[8,79] (FIG. 1f) — a view that offers an intuitive explanation of enhancer specificity. Polymer physics models based on phase separation provided a simple and elegant molecular explanation for the formation of such globules[80–82]: the weak enthalpic interactions driving the phase separation would allow for rapid mixing of chromatin within the globule, facilitating E–P contacts while achieving physical isolation from neighbouring globules (FIG. 2a). Meanwhile, alternative polymer simulations showed that a process of loop extrusion of chromatin by ATP-dependent motor proteins (postulated to be cohesin), which arrest at TAD boundaries, could produce population-average contact behaviours consistent with the Hi-C data despite the contact maps from single cells each looking different to this population average map[83–85] (FIG. 2a). Loop extrusion simulations suggested that TADs are not physical structures (such as globules) that could be seen in single cells the way chromosome territories had been seen[86]. Rather, TADs in these simulations were an emergent statistical property of the ensemble, arising from averaging together many different pairwise contacts that cohesin-loop extrusion created more frequently within TADs than across them. Both models stressed the dynamic nature of the underlying chromatin polymer, which was viewed to be substantially more heterogeneous at the molecular level than classical views of protein folding. These models differed in their outlook of how TADs influence E–P specificity and their expectation of what a snapshot of the 3D structure of a single chromosome might look like at the TAD scale (FIG. 2a).

Whether globular structures exist or the view of a dynamic ensemble of loops was more accurate remained controversial with access only to bulk contact data[87–89]. Results from single-cell Hi-C (scHi-C) experiments provided limited resolution to the controversy, as the data could be interpreted in favour of the existence of globular folds in single cells (supporting the view of TADs as physical structures) but also highlighted the heterogeneity of single-cell data (supporting the view of TADs as statistical features)[90,91]. The ambiguity is understandable. The use of ligation in scHi-C makes clustering analysis more difficult — a particular sequence may be close to multiple others but can only ligate to one, fundamentally limiting the coverage. The understanding of CTCF loop-dots was similarly controversial. Did these represent stable interactions, and was looping of the endpoints required to achieve insulation, as suggested by some genetic experiments[92–94]? Did nested loop-dots represent competing interaction pairs, or cooperative hubs[47,66]? These too, could not be determined with only pairwise average data or sparse scHi-C data.

Additionally, Hi-C data raised new questions about the timing and stability of *cis*-regulatory interactions that were similarly challenging to untangle from bulk, fixed-cell data. For example, it was observed that many genes preferentially interacted with their own enhancers compared with neighbouring ones, either in terms of sharing a TAD or forming a loop-dot, not only in the cell types where the gene was expressed, but also in precursor cells in which the gene was not expressed. To explain why precursor cells exhibit this contact, it was hypothesized that 'pre-formed' topologies might be used in development to poise genes for a more rapid transcriptional response once the cell-type-specific transcription factors were expressed[2,4,95,96] (FIG. 2b). This hypothesis provides an elegant explanation for the role of the structure in both cell types; priming and then firing. It also makes an assumption that is not easily tested with bulk Hi-C assays — that formation of E–P contacts is slow enough and stable enough compared to the other steps of transcription activation that they can be pre-formed and kept in the primed state until the other differentiation signals arise. If E–P contacts are instead quickly formed and quickly lost, they may have no structural and/or functional relevance in the inactive cell type (FIG. 2b). Resolving each of these questions would require a detailed view of the chromatin structure of single cells, not just population-average contact frequencies.

Finally, we have emphasized here some differences in interpretation of the structural basis for E–P communication, along with some of the methodological differences that tend to support one view or the other. However, important common themes have emerged

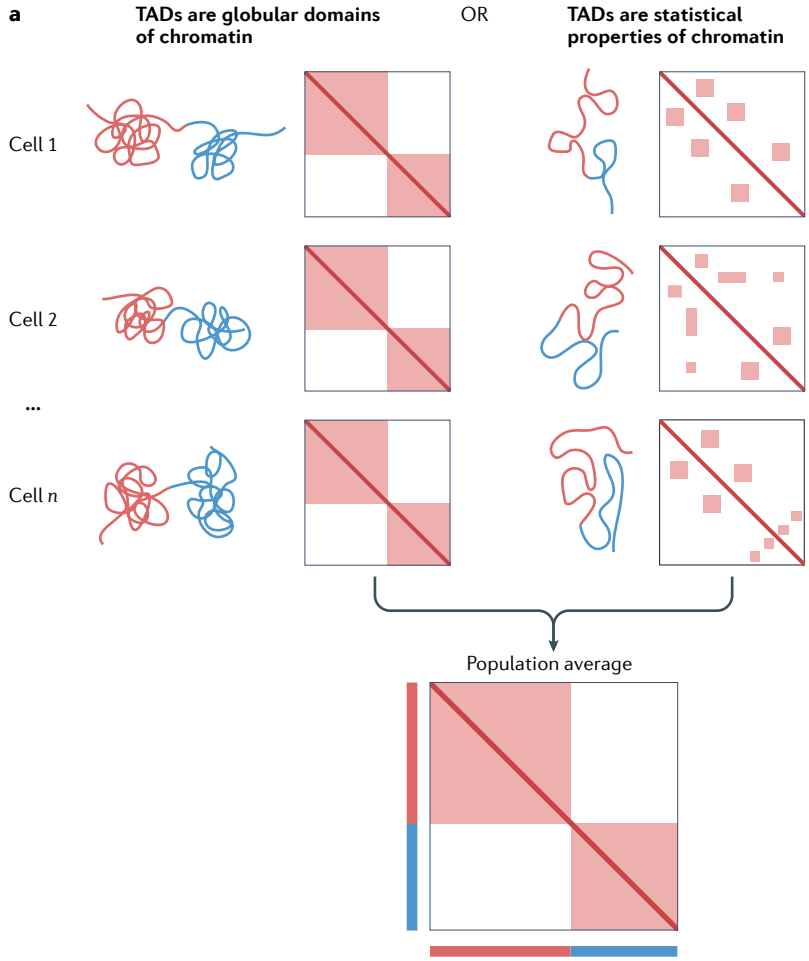


Fig. 2 | **Open questions in the interpretation of population Hi-C data. a** | Alternative models of the single-cell structures underlying topologically associating domains (TADs) observed in Hi-C data. Either TADs reflect phase-separated segments of chromatin, which mix and facilitate contacts between chromatin regions within a TAD but prevent contacts between regions in different TADs (left); or TADs are a purely statistical feature of chromatin, reflecting preferential patterns of pairwise contacts within a generally mixed polymer (right). **b** | Schematics outlining two hypotheses to explain loop-dots, which connect enhancers and target genes in both precursor cell populations in which the gene is not expressed and in differentiated cells in which the gene is actively transcribed. In the first hypothesis, pre-formed loops poise genes for rapid activation upon differentiation, which would require only minor structural changes and/or transcription factor binding. In the second hypothesis, loop-dots reflect dynamic ensembles, not stable structures, whose dynamics do not change upon differentiation. In this scenario, the enhanced frequency of the loop-dot is relevant for inducing transcription only in the differentiated cell.

from these approaches that have strongly shaped the development of the field and we direct interested readers to recent Reviews that provide a comprehensive and historical discussion of the topic[13,16–19].

3D organization under the microscope

It became clear that the field needed new tools with which to visualize the physical 3D structure of the genome in order to better understand the genome, to better test models of the molecular mechanisms that sculpted it and to better understand its variation, both within a cell population and among different cell types. New contrast approaches have enabled electron microscopy to resolve individual chromatin fibers in interphase nuclei[97], but electron microscopy has not provided sequence specificity, nor has it scaled easily to the analysis of large populations of individual cells, both of which are required to address the questions raised by 3C and Hi-C studies. Fluorescence in situ hybridization (FISH) offered sequence-specific labelling, but fluorescence imaging lacks resolution; the wavelength of visible light is orders of magnitude larger than the chromatin fibre, and so diffraction blurs the positions of separate molecules. Into this gap stepped super-resolution fluorescence imaging. Generic super-resolution methods, such as stochastic optical reconstruction microscopy/photo-activated localization microscopy (STORM/PALM), structured illumination microscopy (SIM) or stimulated emission depletion (STED) microscopy (for a review see REFS.[98–100]), combined with new approaches for making synthetic oligonucleotide probes (oligos) (for a review, see REF.[101]), rapidly uncovered previously hidden detail in chromatin structure[48,88,102–105] (also reviewed in REF.[101]). However, these techniques did not robustly resolve individual chromatin fibres in the dense environment of the nucleus (despite their improved resolution compared to diffraction-limited imaging), nor could they distinguish one end of a fibre-like portion of the image from the other (FIG. 3a,b). Thus, these methods lacked the ability to map structures to sequence, and they cannot be used to generate

Diffraction limit
The distance limit below which two objects cannot be distinguished by conventional light microscopy. It depends on the wavelength of light and the objective used.

a Hi-C-like or 3C-like contact map from the image data alone. Similarly, two-colour FISH experiments that label a pair of loci of interest have shown that distance distributions can behave differently from contact frequencies and do not always validate Hi-C results[106]. An approach was therefore needed that could image and distinguish multiple points in the same cell that are positioned at a range of different genomic distances that span the scales of E–P interactions. These criteria have been met by chromosome-tracing super-resolution approaches, which have been enabled by technological advances in rapid sequential hybridization.

Tracing chromosomes with sub-diffraction resolution. In chromosome-tracing approaches, primary oligos bound to the genome carry unique single-stranded DNA barcodes in place of fluorophores and the genomic resolution (in base pairs) of the chromosome trace depends on the spacing between the centres of each barcoded region. Barcodes are detected using a complementary fluorescently tagged secondary oligo, which can be removed after imaging so that their diffracted signal does not overlap with the signal of the next unique secondary oligo. Thus, one step at a time, the DNA path is traced by sequential cycles of fluorophore addition, imaging and removal (FIG. 4a). This process bypasses the diffraction limit to resolution in much the same way that STORM does — detection events are recorded in distinct images separated in time and the centre of each signal can be reliably measured without overlap from adjacent spots. The spatial resolution of the trace (in nanometres) depends on the length of each unique barcode, the signal-to-noise of the spot detection, and the quality of the drift correction (FIG. 4b). Unlike earlier STORM approaches, each spot detected in a trace corresponds to a unique genomic locus, thus allowing reconstruction of the 3D path of the chromatin polymer and enabling mapping of structural features onto the linear genome (FIG. 3c, FIG. 4a). This uniqueness is also beneficial in the identification of false-positive detection events (for example, owing to off-target probe binding) or false-negative events (owing to the failure of some probes to bind in a specific cell) (FIG. 3c). By contrast, generic super-resolution methods for imaging

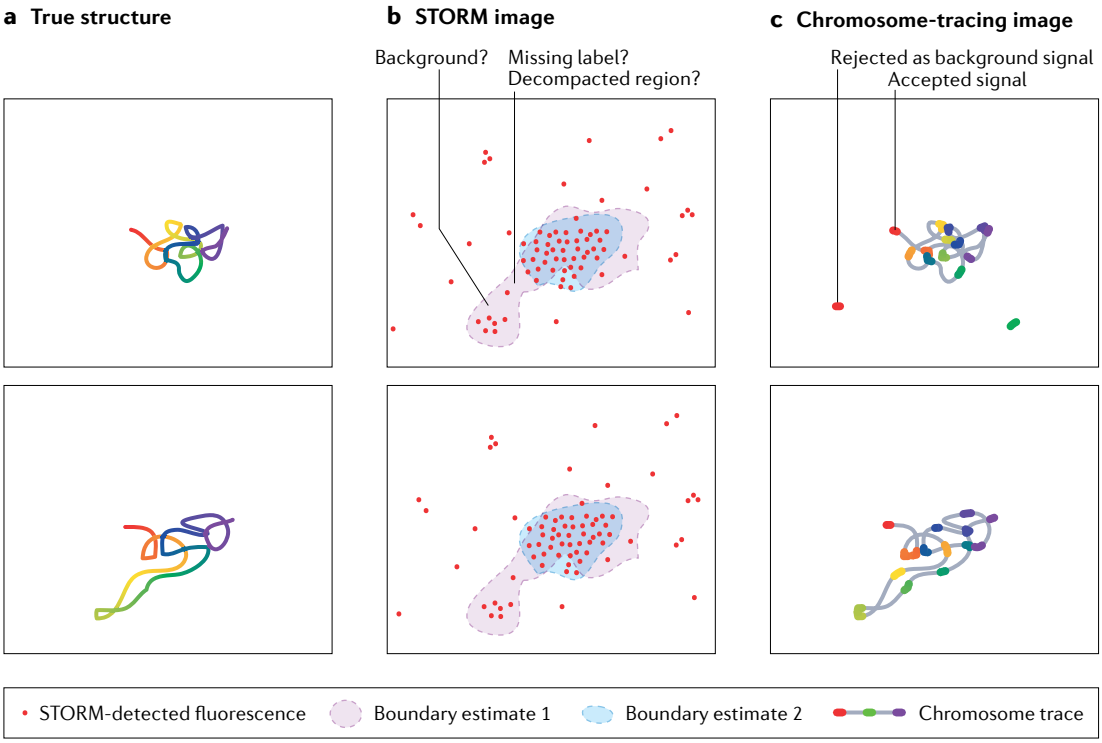
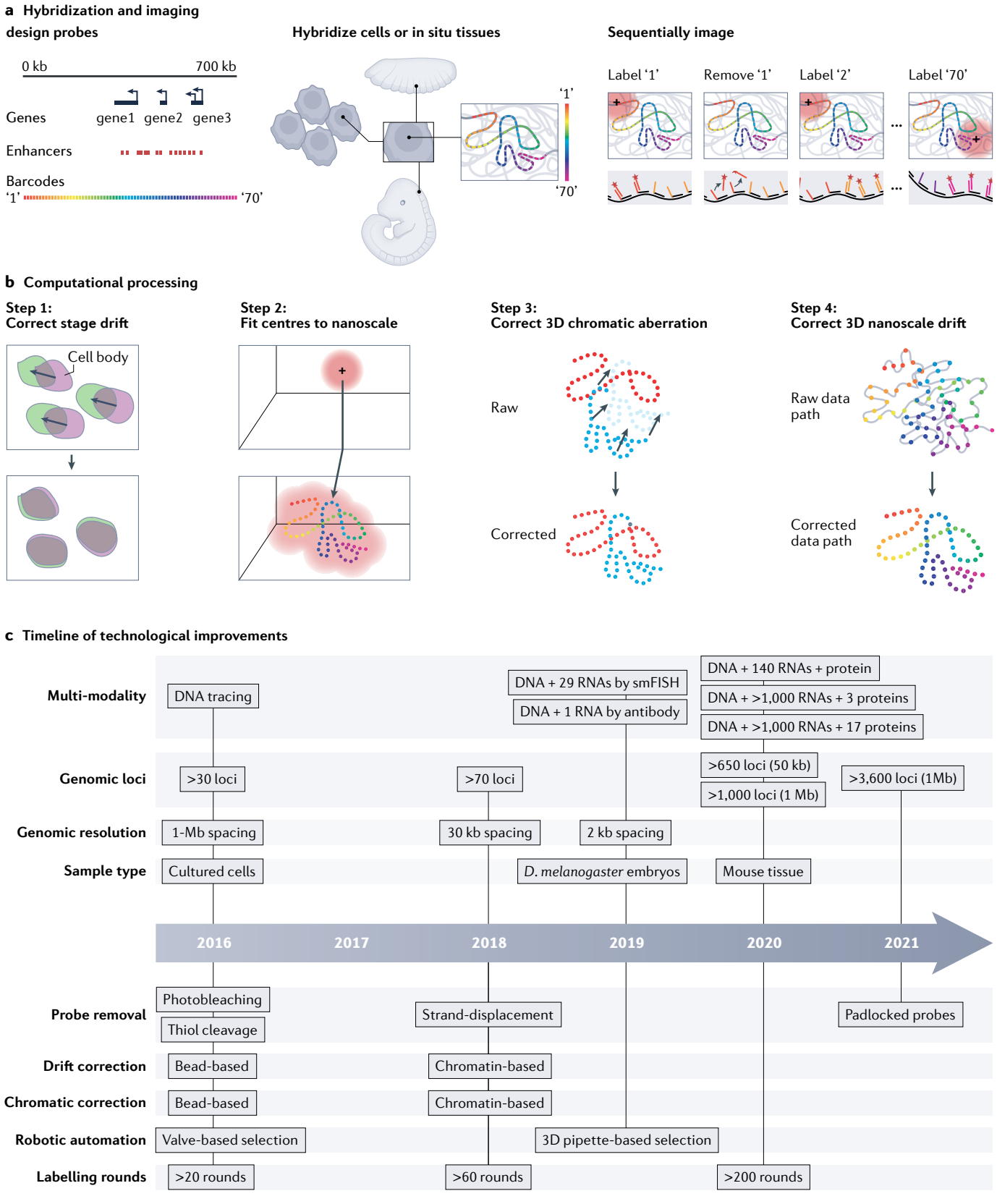


Fig. 3 | **Chromosome tracing disambiguates chromatin structure compared to generic super-resolution approaches.** **a** | Schematic of two different chromatin folds, one compact, and one with an extruded chromatin loop. **b** | Generic super-resolution approaches, such as stochastic optical reconstruction microscopy (STORM, also known as photo-activated localization microscopy (PALM)), generally employ localization-density filters to identify cellular nano-structures, such as chromatin domains, and filter out background localizations (boundary estimates 1 and 2). This approach requires user-defined thresholds, and may filter out decompacted portions of chromatin that loop away from a compact region. The same pattern of detection events may give different estimates of domain volume depending on the threshold used. **c** | In chromosome-tracing approaches, background signal is reduced because on-target signals are generated from an array of fluorescent molecules, not a stochastic blink from a single molecule. Each step along the chromosome trace is denoted with a distinct pseudocolour (see barcode in figure). If two spots are detected in a given step, the one closest to adjacent spots is retained and the more distal spot is rejected as background. This exploits the known connectivity of chromatin polymers (known from the genome sequence) and the known copy number per cell. For cells in which homologous chromosomes are in close proximity, this discrimination becomes more challenging, but can still be aided by underlying knowledge of the polymeric nature of the chromatin, as shown in a recent preprint[162].



chromatin do not as easily distinguish between signal and background (FIG. 3b).

Since its introduction in 2016 (REF.[107]), multiple groups have developed chromosome-tracing methods, including Optical Reconstruction of Chromatin Architecture (ORCA)[108], Hi-M[109], Multiplexed Imaging of Nucleome Architecture (MINA)[110], DNA-MERFISH[111] and DNA-SeqFISH+ (REF.[112]). All these methods attach DNA

◀ Fig. 4 | **Overview of the chromosome-tracing procedure. a** | In situ hybridization decorates the genomic region of interest with oligonucleotide probes, each one appended with a unique synthetic barcode sequence. By sequentially hybridizing, imaging and then removing fluorescent oligos complementary to these barcodes, the chromosome trajectory is traced. Because the individual steps are imaged sequentially, the diffractive signals of the fluorescent oligos do not overlap to blur the structure. **b** | Computational processing reconstructs the chromosome path from the imaged oligos. In step 1, the stage drift between sequential hybridization rounds is corrected by rigid translation. Beads, cell bodies, or fluorescent chromatin tags can be used as fiducial markers for this step. In step 2, the 3D centroid of the diffraction-limited signal detected in each image is computed, typically using a 3D Gaussian fit or radial symmetry approximation. In step 3, if multiple fluorescent channels are used per hybridization round, the chromatic aberrations between channels are corrected. This process requires multicoloured fluorescent markers distributed in 3D throughout the field of view. Multicoloured beads attached to the surface of the cells and the cover glass can work, although they poorly approximate the aberrations deeper in the sample away from the coverglass. Multicoloured tags targeted to the same DNA locus typically cover the 3D field better. These data are used to compute the chromatic correction map. Finally, in step 4, cell- or chromosome-specific nanoscopic drift is corrected using fiducial labels attached to the chromosome. The resulting registered, chromatically corrected, drift-corrected coordinate list tracks the 3D structure of the chromosome. Replicate labels of a subset of barcodes provide a final measure of the precision of the reconstruction. References [114–116] are recommended for a detailed description of the procedures. **c** | A timeline of notable advances in chromosome-tracing approaches.

barcodes to distinct genomic loci — sometimes also to distinct RNA targets and even protein targets — and use sequential hybridization and removal of secondary oligos to distinguish the targets. These parallel efforts have led to many improvements: in genomic resolution (from 1-Mb resolution[107] to around 2-kb resolution[108]); in genomic throughput (from the low tens of signals in 2016 (REF.[107]) to over 600 unique signals in 2020 (REF.[111])); in the use of combinatorial signals (allowing 1,000–2,000 loci to be probed at 1-Mb resolution)[112,113]; in the number of cells that can be analysed in a single experiment (from hundreds[107,112] to over 100,000 (REF.[108])); and in multimodal imaging that combines chromosome tracing with detection of RNA, nascent RNA, and/or protein and enables analysis in parallel distinct cell types from the same tissue in situ[108–112] (FIG. 4c).

Several noteworthy innovations have supported these steady improvements (FIG. 4c) (for more detail, see REFS.[114–116]). New chemical approaches for removing fluorescent labels between imaging rounds have accelerated data collection by improving the speed with which old signals are removed. The first approach relied on photobleaching, which is slow. The use of fluorophores, which could be chemically cleaved from the oligos and applied to all cells at once, accelerated the process but precluded replicate labelling[117]. Subsequent work used circularized padlock probes to retain the primary probe on the sample while denaturation stripped the secondary probe[112,113,118]. The fastest approach, which is also fully reversible, uses strand-displacement oligos to peel off the last-imaged secondary oligo in the same reaction that the next secondary oligo anneals[108,116,119]. The experimental setup has been simplified and run costs have been reduced by implementing robotic automation that is compatible with 96-well plates and lower reagent volumes[108,116] and by the use of universal fluors that can be connected to unique primary barcodes by bridge-oligos[108,116,119,120]. Faster, simpler data collection has contributed substantially to an increase in the typical number of loci studied in each experiment over the years. Recently, multiplexed barcoding approaches introduced by the Zhuang and Cai laboratories[111,112] have enabled thousands of spatially distributed target loci to be identified in a single cell. Rather than using thousands of sequential imaging rounds to identify unique targets, these approaches image distinct combinations of loci in each round such that the combination of times a target appears is unique. However, combinatorial approaches largely forfeit sub-diffraction resolution: in order to decode the multiplexed label, it is important that targets do not overlap, which requires targets to be a certain distance apart (generally >1 Mb) in the genome, thereby reducing the genomic resolution achievable compared to using unique labels. Although notable limitations remain for chromosome-tracing approaches (BOX 2), we can expect considerable further advances in the near future, given the rapid proliferation of and improvements in these methods. Indeed, these chromosome-tracing experiments have already started to provide answers to some of the unresolved questions arising from sequencing-based studies.

Fiducial markers
Markers used to correct for drift that may occur during an experiment. These can be fluorescent beads or labels on the DNA that remain constant throughout the imaging experiment.

Enhancer–promoter communication: loop or domain?

Sequencing-dependent approaches suggested two alternative views of E–P communication: point-to-point organization in sequence-specific loops; or domain organization, such as local confinement in TADs (FIG. 1c,f). At the population average, chromosome-tracing experiments consistently revealed patterns of frequent proximity that strongly resemble the domain structure seen in Hi-C data, rather than a network of specific E–P loops. The first example was the large gene desert upstream of the gene encoding the metalloproteinase ADAMTS1; in human fibroblasts, the frequency-of-proximity (also called contact-frequency) profile mapped by chromosome tracing revealed block-like domains with sharp transitions to lower contact-frequency regions, clearly matching the TAD pattern seen in deep Hi-C data[47,119] (FIG. 5). Loop-dots at CTCF sites emerge from the average chromosome tracing, much as in the Hi-C data, further supporting the conclusion that these are not accessibility artefacts[47,119] (FIG. 5). Results from subsequent chromosome-tracing experiments from several teams were similarly consistent with Hi-C data, showing primarily block-like rather than peak-like organization of the genome across diverse samples, including *D. melanogaster* embryos[108,109], mouse liver[110], mouse brain[113], and a substantial number of cultured cell types[64,108,111,112,119–121]. High-throughput imaging, using 125 pairs of bacterial artificial chromosome-based probes rather than chromosome tracing, provided further agreement with Hi-C data[122]. Thus, imaging-based approaches support the view that the importance of enhancers and promoters in autonomously controlling their interactions had been overstated by 3C assays, and that additional factors instead drive a domain-type, rather than point-to-point type, structure.

By combining RNA labelling with chromosome tracing, molecular and spatial signals can be used to distinguish cell types in intact tissues, thereby enabling the study of cell-type-specific enhancer regulation in its in vivo,

Box 2 | **Limitations of microscopy approaches for studying 3D organization of transcription control**

Microscopy approaches to the study of nanoscale chromatin organization and transcription control are subject to numerous sources of potential error, bias and uncertainty. Moreover, these limitations are much less well understood than comparable limitations in sequencing-based approaches, such as bias in enzymatic activity, sequence recovery and amplification. Here we discuss a non-comprehensive list of a few notable limitations of microscopy-based approaches, and highlight some recent work to quantify and minimize some of these sources of error.

Fixation artefacts. In principle, chemically freezing the sample to get a snapshot view is not problematic. However, artefacts can be introduced if the sample is fixed too slowly and stress pathways activate, or too little so that only more reactive molecular species are immobilized or are immobilized before the less reactive ones. A substantial range of fixation approaches exist, some of which are sensitive to subtle changes in execution, and the chosen approach should preserve the features of interest without interfering with subsequent steps in the analysis. There is no need to preserve atomic-scale protein folding if that folding is well below the resolution limit of the assay, nor to preserve 3D positioning down to nanometre scale if the resolution limit is in the tens of nanometres.

Labelling artefacts. Labelling is likely to be more disruptive than the preceding fixation step, as hybridization of oligo probes requires denaturing of the DNA, which disrupts the double-helix structure and probably also nucleosome wrapping. Denaturants will also unfold some proteins. However, all these changes can be accomplished with molecular-scale movements too small to be detected in methods with 10–30 nm resolution and, provided that the fixation creates a strong enough and dense enough gel out of the cellular proteins to prevent larger-scale motion, DNA and protein unfolding per se are not resolution-limiting to current approaches. Instead, resolution limits are dominated by the precision of probe localization and the size of the probe or probed region of chromatin (see below). Notably, new approaches, including CasFISH[168], GoldFISH[169], Raser-FISH[170] and emerging new applications of Raser-FISH[171], bypass the effects of global denaturing by using enzymes such as Cas9, helicases and exonucleases to generate localized single-stranded targets. Although these alternative approaches are sorely needed, we recommend that enzymatic sequence biases and other artefacts that perturb the structure be investigated with chromosome tracing and orthogonal approaches such as in situ Hi-C. It should be noted that existing denaturation methods based on heat and acid show excellent agreement with gold-standard deep Hi-C datasets, which depend on alternative matrix balancing adjustments for detection bias. Whether or not these new methods allow improved resolution of TADs, loop-dots and E–P interactions, or whether enzymatic sequence bias and perturbations of labels introduced into pre-fixed cells abrogate these, remains to be tested. Meanwhile, their improved compatibility with protein labelling is sure to facilitate new studies.

Drift and degradation. For approaches such as chromosome tracing that reconstruct an image by repeated probing of the same sample over hours to weeks, errors in the repositioning accuracy of the hardware, and motion of the tissue, individual nuclei and individual chromosomes under thermal fluctuations or degradation potentially confound the spatial relationship between the probes. Fortunately, notable improvements have been made in recent years to measure these errors, although adoption of such controls is far from complete. For example, development of strand-displacement probes allows the same element to be re-probed multiple times and throughout the experiment, providing a direct measure of how much uncorrected motion has occurred[108,116,119], which has prompted development of sample preservation approaches for long experiments[108,111]. Currently reported uncertainties from these optimized methods are smaller (20–50 nm) than the typical spacing between adjacent probe regions (100–300 nm, depending on probe resolution)[108,119].

Label size. Label size is often a chief, but sometimes underappreciated, contributor to limited resolution. The ability of microscopy approaches to measure spatial organization in units of nanometres, rather than relative read counts, has invited comparisons to the physical dimensions of other components of the transcription machinery, such as the Mediator complex or holoenzyme. These comparisons have led some to challenge long-standing models of enhancer-mediated gene regulation, suggesting that enhancer-bound factors can regulate transcription even though hundreds of nanometres separate them from their target TSSs[14,18,63,172,173]. A potential difficulty with this conclusion is that the labels themselves are not atomic-sized elements, but instead short, flexible stretches of chromatin frequently 100–300 nm long, as revealed by chromosome tracing in fixed cells with resolutions of 2–3 kb (REFS.[108,133]) or 5 kb (REFS.[64,108,110]), and by recent live imaging of adjacent probes[137,138]. Consequently, molecules bound within one labelled region may be in molecular contact with molecules bound to another, even though the fluorescent centres of each region are measured to be several hundred nanometres apart. This problem is compounded in live imaging if larger (10 kb) probes are used to enhance contrast and they are positioned a few kilobases distal to the enhancer or promoter to avoid disrupting it, and the motion blur of live imaging is added.

Availability. In contrast to sequencing-based methods, chromosome tracing and related multiplexed imaging technologies are currently less readily accessible to the broad research community. Most work has been performed on home-built optical instruments, with home-built fluidic systems, and analysed by home-built software environments, all of which require notable expertise to maintain and use efficiently. Extensive sharing of detailed protocols and software have facilitated some broader adoption of these methods[114–116], but they are still far from achieving the availability that commercial instruments and kits make possible for sequencing assays.

multicellular context. For example, in *D. melanogaster* embryos the enhancers of the *Ubx*, *abd-A*, *Abd-B*[108] and *snail*[109] genes reside within contact domains (that is, TADs) with their promoters, and E–P contacts are no more frequent than contacts among non-regulatory elements in the same domain. The contact domains around the *D. melanogaster hox* genes are particularly noteworthy for their striking alignment to functional enhancer domains mapped by random integration of minimal reporter genes throughout this locus[123] (FIG. 6a). Each functional enhancer domain corresponds to a distinct contact domain. The boundaries of the contact domains change in different cell types so that active enhancer and active repressor elements are always clustered with the promoters they control (FIG. 6b). Confining E–P interactions to a domain, rather than an E–P specific bridge, also helps to explain how a randomly inserted reporter gene that carries a heterologous promoter can be activated by any of the enhancers in its domain[52] (FIG. 6a) Collectively, these in vivo microscopy experiments further support the domain-like organization of *cis*-regulatory elements and their target genes rather than a model involving E–P loops.

TADs are heterogeneous ensembles, not globules. Although the aggregate frequency data from chromosome tracing supported the domain model of E–P interaction, the physical structures observed in single cells contrasted with the popular depiction of TADs as globules (FIGS. 1f,2a). In all the published chromosome-tracing

Entropic
Changes that increase the number of accessible microstates in the system and do not require input energy.

data, the chromatin corresponding to the TAD adopts heterogeneous, often intermixed folds, rather than collapsed, isolated structures (FIG. 6c). The preference for intradomain contacts over interdomain contacts is only fully apparent when the whole ensemble is considered (for example, mean or median domain distances). Thus, rather than denoting globular structures that cluster and insulate E–P interactions, TADs are emergent statistical patterns that arise from the flexible, probably dynamic, motion of the chromatin polymer, which is captured in a different state in each cell (FIG. 2a). Individual chromatin traces may nonetheless exhibit frequent globular folds. These have been called 'TAD-like' domains[119] or 'chromatin nanodomains'[103] (FIG. 6c). Notably, the number and boundaries of such domains are variable cell-to-cell and individually distinct from the population average[103,119]. As such, they cannot be equated with TADs. Rather, they reflect the generally globular nature of chromatin, which polymer physics demonstrates to be characteristic of any polymer experiencing dense confinement and/or preferential self-interactions compared to solvent interactions[124–127].

This distinction between physical structure and emergent statistical pattern has important consequences for our understanding of how TADs affect E–P interactions and transcription. Many theories on the behaviour of TADs had been inspired by comparison to chromosome territories, which similarly appear as block-like domains in a Hi-C map[47,79,89,128]. However, it is possible for a particular promoter to reside 'inside' or 'near the surface of' a chromosome territory, because these territories are physical structures readily identifiable within a single cell, as shown by a number of elegant microscopy studies[86,129,130]. By contrast, it is not possible for a promoter to reside 'near the surface of the TAD', as the TAD is a statistical property of the whole population, not a structure as such within an individual nucleus. Notably, this view of TADs as properties of structural ensembles rather than structures themselves is consistent with the earlier coarse-grained mechanistic models from polymer physics, such as loop extrusion[127,131]. However, it can also be reconciled with degenerate phase separation[132], whereby preferential interactions between distinct chromatin types drive interaction frequency without invoking the need for ATP-dependent motor proteins[132]. As such, this heterogeneity does not distinguish to what extent, if any, energy-expending processes (such as loop extrusion) or entropic processes (such as phase separation) each contribute to shaping the ensemble of 3D structures.

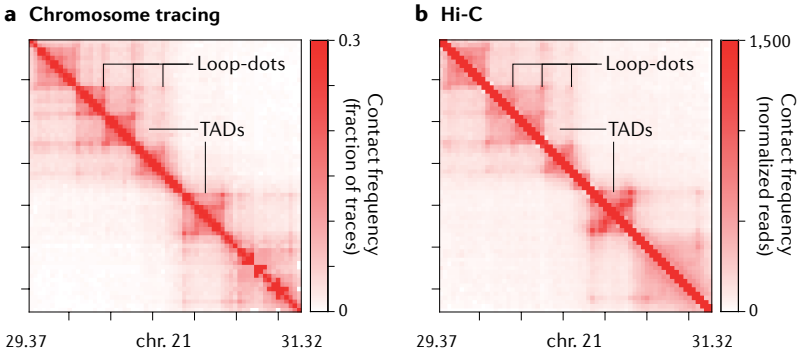
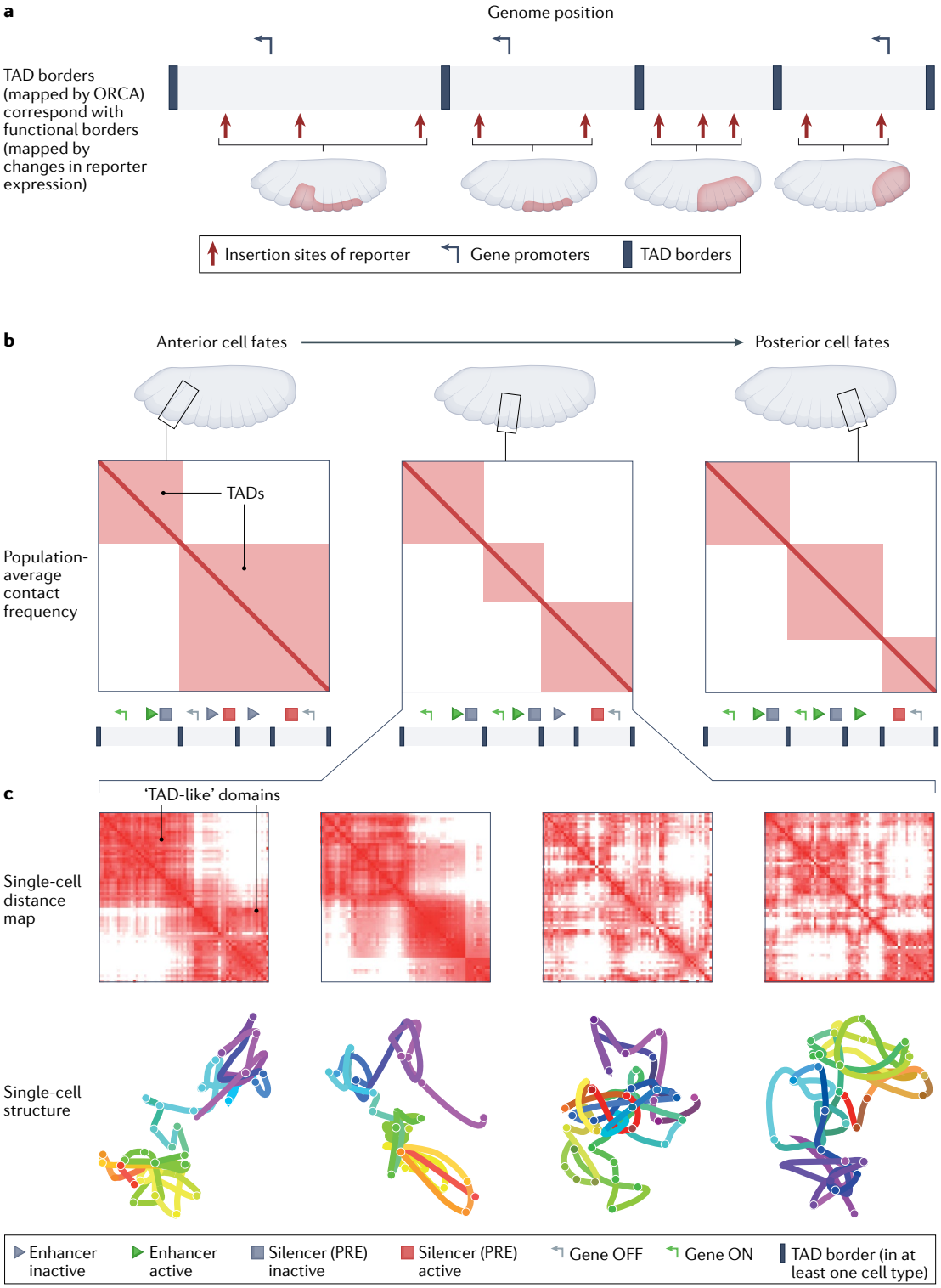


Fig. 5 | **Comparison of chromosome-tracing data and Hi-C data. a** | A contact-frequency map generated from chromosome-tracing data on a human fibroblast cell line (IMR90), computed using a 150-nm contact threshold on over 3,000 traces. Trace data was taken from REF.[119]. **b** | A contact-frequency map generated from Hi-C data on the same cell line and same genomic region, showing normalized read counts after matrix balancing. Data plotted taken from REF.[47] (GSE63525). A few examples of loop-dots and TADs are indicated in each map.

TADs and loops emerge from chromatin dynamics. Although imaged in fixed cells, these heterogeneous structures seen among the chromosomes of different cells are widely interpreted as evidence of a dynamic ensemble, whereby a single chromosome moves through the different states seen in the population[108,127,133,134]. If instead each cell retained, throughout interphase, the highly unique fold in which it was captured upon fixation, it is difficult to imagine how nanoscale 3D organization and TADs could affect transcriptional control: in one cell, chromatin from two different TADs with distinct enhancers and promoters may be well mixed together, in another, they may be well separated. It is not until hundreds to thousands of such diverse individual 3D structures are averaged together that the TAD patterns arise. If the relative frequencies of each structure across the population are assumed to reflect the relative amount of time an individual cell spends in each state over interphase, the individual cell will, in a temporal statistical sense, experience a TAD border, thus giving an intuitive explanation of how TADs influence transcription. Rapidly advancing live-imaging experiments have started to provide estimates of how quickly individual cells might transition through these different 3D structures by measuring temporal dynamics of individual pairs of points[63,135–138]. In the meantime, fixed-cell experiments can bound the timescales of TAD-scale dynamics. In chromosome-tracing experiments, TADs emerge in the population statistics within hours of mitosis in G1 cells, look similar in G2, and are rapidly lost upon cohesin removal[119]. More detailed estimates of these formation times have been obtained from Hi-C experiments after cohesin removal or mitotic exit[139–141].

The view of chromatin organization as a dynamic ensemble of contacts, rather than a stable structure, has also affected discussions of the developmental timing of E–P contacts, particularly the role of preformed structures (FIG. 2b). Investigation of the *cis* regulation of the *docs* genes in *D. melanogaster* with a chromosome-tracing approach revealed high-frequency contacts among the promoters of the three *docs* genes and their enhancers[133], seen as loop-dots in the population average data. These loop-dots arose before gene activation, and persisted in cell populations regardless of whether *docs* was active or silent, as seen with other pre-formed loops or TADs[2,4]. Yet, as for TADs in earlier imaging studies, the enhancers and promoters engaged in these loop-dots were in proximity in only 10–30% of cells, suggesting that the interactions are dynamic and transient[133]. Similarly, the loop-dot connecting the *Sox2* promoter and its embryonic stem cell enhancer was found by chromosome-tracing experiments to represent a contact in less than 30% of mouse embryonic stem cells[64,108]. Imaging of human fibroblasts showed



that the CTCF loop-dot anchors along the domains, which are separated by a few hundred kilobases to over a megabase, also occur in less than 30% of cells (FIG. 5), arguing that these too are transient interactions[119]. Whereas these absolute frequencies are somewhat dependent on the contact threshold chosen, the observation that loop-dots are transient features remains robust because, in contrast to other DNA FISH-based approaches, tracing approaches map the loop-dots and can determine the range of contact distances wherein the loop-dots and TADs exist. Thus, the emerging microscopy data undermines the hypothesis of pre-formed loops for rapid activation, instead suggesting that even loops denoted as stable between cell types are constantly reforming and dissolving at the single-cell level. The relative stability or frequency of loops may be cell-type-dependent (as in the *Hox* gene example) or invariant (as with *docs*), but all interactions are transient (FIG. 2b). Further live-cell

◀ Fig. 6 | **Cell-type-specific topologically associating domains link cell-type-specific combinations of regulatory elements in *Drosophila melanogaster* embryos.** **a** | Schematic depiction of the agreement between spatial domain boundaries identified by chromosome tracing and the functional domains mapped by reporter-gene insertions. We note that reporter expression patterns are shared by all insertions within a spatial domain and are distinct across spatial domain borders. The illustration is schematic and does not reflect precise genome coordinates or expression patterns. **b** | Transcription of the Hox genes activates in a segment-specific manner in the *D. melanogaster* embryo. In the most anterior segments, gene activation coincides with the separation of the promoter and its enhancer into a distinct topologically associating domain (TAD), distinct from neighbouring repressed genes and silencers (left). In segments towards the middle of the embryo, adjacent active genes responding to distinct enhancers separate into distinct TADs (middle). In more posterior segments, some TADs expand to include additional enhancers (right). Promoters that share a TAD with active silencers are inactive. Regulatory element spacing is schematic. **c** | Illustration of single-cell distance maps (top) and corresponding chromosome traces (bottom). We note the existence of box-like regions in some of the single-cell distance maps that resemble the box-like appearance of TADs in the population average data, but which vary in their genomic coordinates cell-to-cell. We note also the considerable variation in 3D structure within a cell type, which together average to a reproducible, cell-type-specific set of TADs. PRE, polycomb response element.

measurements will be necessary to determine the rate at which loops form and dissolve within the population, and recent experiments have estimated interactions on the scale of 15 minutes for two different CTCF-loops in mouse embryonic stem cells[137,138].

Future perspectives

With these advances in visualizing the 3D structure of chromatin, resolved at the length scales relevant for *cis*-regulatory transcriptional control, several long-standing questions in transcriptional regulation may soon be addressed. Here, we highlight a few areas in which we hope to gain insights in the near future.

How do multiple enhancers regulate the same gene? Genetic experiments make it clear that *cis*-regulation is inherently a more-than-pairwise affair, with many promoters integrating signals from multiple enhancers and silencers in a given cell type, and many enhancers influencing the expression of more than one promoter[1]. To what extent these interactions are structurally competitive, cooperative or independent is poorly understood. When do enhancers form 3D regulatory hubs and when do they displace one another for control? The investigation of these inherently multiway interactions has just begun to be tackled[142–145], and we expect that the ability of chromosome-tracing methods to visualize and distinguish all enhancers, promoters and intervening sequences at once, without concern for ligation or digestion bias, and to superimpose the spatial organization of RNA and protein components on these structures will continue to facilitate these investigations. Both sequencing and microscopy approaches have shown that epigenetic regulators of transcription, such as Polycomb group proteins, HP1 group proteins, and Trithorax group proteins, influence the 3D organization of the genome[146–148]. Members of all these protein groups have also been shown to form phase-separated condensates[149–152]. Understanding the physical nature of this organization at the single-cell level and its potential functional importance to gene activation or repression will also be aided by visualizing how they affect chromosome traces.

How can transient, heterogeneous contacts enable precise gene expression? From chromosome-tracing studies on fixed cells, and now increasingly with live imaging, we have seen that 3D genome folding is heterogeneous and highly dynamic. Fixed-cell experiments that combine chromosome tracing with RNA labelling in the same cells have shown statistically significant[108], yet weak, correlation between E–P contact frequency and nascent RNA levels. In several recent works, the probability of observing nascent transcription from a promoter with the enhancer in proximity was only slightly higher (0–30%) than for promoters where the enhancer was hundreds of nanometres distal[64,108,153]. This conclusion largely paralleled that from live imaging of the *Sox2* E–P pair, for which no statistically significant correlation was detected between enhancer and promoter proximal tags and bursts of nascent RNA[63].

The limited correlation between transcription and E–P distances in single cells implies two non-exclusive conclusions about transcriptional regulation by distal enhancers: either the 3D structure changes rapidly in the few minutes it takes polymerase to transcribe the first few kilobases of the gene (from which nascent RNA production is detected); or individual enhancer contacts are not triggering events for transcriptional bursts, and other biochemical processes at the promoter may enable it to count the frequency of interactions, rather than simply fire each time an active enhancer is encountered. Two recent biochemical models have described in theory how such counting might work[154,155]. More provocatively, these analyses have shown that such mechanisms can also explain the nonlinear relationship between the average expression level of a gene and the average contact frequency of its enhancers[154,155]. Another possibility is that there is more information in genome folding beyond E–P distances. Convolutional neural networks found that the power of 3D organization to predict nascent transcription state using the entire *cis*-regulatory domain was substantially improved relative to using E–P proximity alone, although it was still only modestly predictive at a cell-to-cell level[156]. To understand how the heterogeneity in DNA folding is still informative for transcription, further use of both fixed-cell and live-cell imaging, combined with modelling and experiments to test the model predictions, will be required. It will also be necessary to study a larger number of different E–P interactions to understand the breadth of mechanisms, given that some genes might show a much stronger correlation between E–P proximity and promoter firing than others, as shown with three-colour live imaging for a synthetic E–P contact in *D. melanogaster*[136].

What structural features are relevant for gene regulation? As seen in the recent analyses of pre-formed loops, some structural features of the genome might not have a clear (or possibly any) function in all of the cell types in which the feature is present, which does not preclude the possibility that the structure is important for transcription in some cell type and/or environmental condition. Thus, some recent discussions about the potential irrelevance of 3D genome organization to transcription[53,157],

Convolutional neural networks
(CNNs). Algorithms designed to learn from the data to uncover connections. CNNs are frequently used in image recognition and have been increasingly used to uncover relationships in biological data.

based on repeated observations in which TAD or loop-dot perturbation fails to measurably affect transcription, may be premature. Future analyses in vivo, across more diverse cell types and more perturbations, will help us to understand better when 3D organization is, and when it is not, of regulatory significance. The continued partnership of sequencing approaches, with their ability to rapidly profile the whole genome at high resolution, and chromosome-tracing approaches, with their ability to measure single-chromosome folding, will be essential. A broad collection of emerging microscopy-dependent methods beyond chromosome tracing[10,13,158–161], which we have not reviewed here, will add further dimensions to these analyses.

Conclusions

Chromosome-tracing microscopy technologies have allowed 3D chromatin structures to be visualized across many thousands of single cells. These data largely support the view that E–P interactions are regulated by contact domains, as opposed to the view that enhancer and promoter elements largely direct their own interactions. Moreover, imaging has contributed clear evidence to support the view that TADs are emergent statistical features of chromatin, rather than structural building blocks such as the α-helix, chromosome territory or nucleus. Other prominent features first identified by Hi-C, such as the loop-dot interactions between CTCF sites that demarcate the borders of many TADs, were also revealed by imaging to be relatively rare in any fixed snapshot of the cell population. The sparsity of these loops at the single-cell level, and the heterogeneity of structures in single chromosomes that averaged to TADs, strongly suggested a dynamic 3D structure and provided an estimate of the relative frequencies of the different underlying configurations that average over time to make a TAD or loop-dot. The agreement between population-level Hi-C data and microscopy measurements not only gives confidence in the data but has proved complementary in providing answers to some of the outstanding questions in gene regulation, as described above. However, the examples published so far have only scratched the surface of how the 3D structure of the genome contributes to the regulation of transcription. In the next few years we can expect that combining the power of 3C methods, FISH omics, live imaging of chromatin and modelling will bring further important discoveries and surprises to the field.

Published online 14 September 2022

1. Levine, M., Cattoglio, C. & Tjian, R. Looping back to leap forward: transcription enters a new era. *Cell* **157**, 13–25 (2014).
2. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
This review discusses enhancers, transcriptional regulation and the role of 3D genome organization, focusing on models and hypotheses.
3. Robson, M. I., Ringel, A. R. & Mundlos, S. Regulatory landscaping: how enhancer-promoter communication is sculpted in 3D. *Mol. Cell* **74**, 1110–1122 (2019).
4. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
5. Marinić, M., Aktas, T., Ruf, S. & Spitz, F. An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**, 530–542 (2013).
6. Lorberbaum, D. S. et al. An ancient yet flexible *cis*-regulatory architecture allows localized Hedgehog tuning by *patched/Ptch1*. *eLife* **5**, e13550 (2016).
7. Ibrahim, D. M. & Mundlos, S. Three-dimensional chromatin in disease: what holds us together and what drives us apart? *Curr. Opin. Cell Biol.* **64**, 1–9 (2020).
8. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170–1187 (2016).
9. Santiago-Algarra, D., Dao, L. T. M., Pradel, L., España, A. & Spicuglia, S. Recent advances in high-throughput approaches to dissect enhancer function. *F1000Res.* **6**, 939 (2017).
10. Agbleke, A. A. et al. Advances in chromatin and chromosome research: perspectives from multiple fields. *Mol. Cell* **79**, 881–901 (2020).
11. Shaban, H. A., Barth, R. & Bystricky, K. Navigating the crowd: visualizing coordination between genome dynamics, structure, and transcription. *Genome Biol.* **21**, 278 (2020).
12. Shaban, H. A. & Seeber, A. Monitoring the spatio-temporal organization and dynamics of the genome. *Nucleic Acids Res.* **48**, 3423–3434 (2020).
13. Jerkovic´, I. & Cavalli, G. Understanding 3D genome organization by multidisciplinary methods. *Nat. Rev. Mol. Cell Biol.* **22**, 511–528 (2021).
This relatively comprehensive article reviews the primary methods used in studying chromatin structure and organization.
14. Lim, B. & Levine, M. S. Enhancer–promoter communication: hubs or loops? *Curr. Opin. Genet. Dev.* **67**, 5–9 (2020).
15. Dekker, J. et al. The 4D nucleome project. *Nature* **549**, 219–226 (2017).
16. Goel, V. Y. & Hansen, A. S. The macro and micro of chromosome conformation capture. *Wiley Interdisc. Rev. Dev. Biol.* **10**, e395 (2021).
17. Mirny, L. & Dekker, J. Mechanisms of chromosome folding and nuclear organization: their interplay and open questions. *Cold Spring Harb. Perspect. Biol.* **14**, a040147 (2021).
This recent review summarizes major advances in the field (particularly from Hi-C and modelling), and highlights many open questions for the field.
18. McCord, R. P., Kaplan, N. & Giorgetti, L. Chromosome conformation capture and beyond: toward an integrative view of chromosome structure and function. *Mol. Cell* **77**, 688–708 (2020).
19. Dekker, J. & Mirny, L. A. The 3D genome as moderator of chromosomal communication. *Cell* **164**, 1110–1121 (2016).
20. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a B-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308 (1981).
21. Banerji, J., Olson, L. & Schaffner, W. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **33**, 729–740 (1983).
22. Driever, W., Thoma, G. & Nüsslein-Volhard, C. Determination of spatial domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the bicoid morphogen. *Nature* **340**, 363–367 (1989).
23. Stanojevic, D., Small, S. & Levine, M. Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. *Science* **254**, 1385–1387 (1991).
24. Chung, J. H., Whiteley, M. & Felsenfeld, G. A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74**, 505–514 (1993).
25. Udvardy, A., Maine, E. & Schedl, P. The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *J. Mol. Biol.* **185**, 341–358 (1985).
26. Kellum, R. & Schedl, P. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**, 941–950 (1991).
27. Dekker, J. & Misteli, T. Long-range chromatin interactions. *Cold Spring Harb. Perspect. Biol.* **7**, a019356 (2015).
28. Han, J., Zhang, Z. & Wang, K. 3C and 3C-based techniques: the powerful tools for spatial genome organization deciphering. *Mol. Cytogenet.* **11**, 21 (2018).
29. de Wit, E. & de Laat, W. A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* **26**, 11–24 (2012).
30. Sati, S. & Cavalli, G. Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* **126**, 33–44 (2017).
31. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
32. Deng, W. et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
33. Palstra, R.-J. et al. The β-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* **35**, 190–194 (2003).
34. Kagey, M. H. et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
35. Amano, T. et al. Chromosomal dynamics at the *Shh* locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* **16**, 47–57 (2009).
36. Noordermeer, D. et al. The dynamic architecture of *Hox* gene clusters. *Science* **334**, 222–225 (2011).
37. Montavon, T. et al. A regulatory archipelago controls hox genes transcription in digits. *Cell* **147**, 1132–1145 (2011).
38. Schoenfelder, S. & Fraser, P. Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
39. Phillips-Cremins, J. E. et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **4**, 3–7 (2013).
40. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
41. Mumbach, M. R. et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat. Genet.* **49**, 1602–1612 (2017).
42. Li, G. et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84–98 (2012).
43. Fullwood, M. J. et al. An oestrogen-receptor-α-bound human chromatin interactome. *Nature* **462**, 58–64 (2009).

44. Fang, R. et al. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.* **26**, 1345–1348 (2016).
45. Lieberman Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
46. Imakaev, M. et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* **9**, 999–1003 (2012).
47. Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
48. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
49. Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
50. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
51. Le Dily, F. et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* **28**, 2151–2162 (2014).
52. Symmons, O. et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **24**, 390–400 (2014).
53. Cavalheiro, G. R., Pollex, T. & Furlong, E. E. To loop or not to loop: what is the role of TADs in enhancer function and gene regulation? *Curr. Opin. Genet. Dev.* **67**, 119–129 (2021).
54. Ibrahim, D. M. & Mundlos, S. The role of 3D chromatin domains in gene regulation: a multi-facetted view on genome organization. *Curr. Opin. Genet. Dev.* **61**, 1–8 (2020).
55. Paliou, C. et al. Preformed chromatin topology assists transcriptional robustness of *Shh* during limb development. *Proc. Natl Acad. Sci. USA* **116**, 12390–12399 (2019).
56. Williamson, I. et al. Developmentally regulated *Shh* expression is robust to TAD perturbations. *Development* **146**, dev179523 (2019).
57. Williamson, I., Lettice, L. A., Hill, R. E. & Bickmore, W. A. *Shh* and ZRS enhancer colocalisation is specific to the zone of polarising activity. *Development* **143**, 2994–3001 (2016).
58. Symmons, O. et al. The *Shh* topological domain facilitates the action of remote enhancers by reducing the effects of genomic distances. *Dev. Cell* **39**, 529–543 (2016).
59. Lupiáñez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene–enhancer interactions. *Cell* **161**, 1012–1025 (2015).
60. Kraft, K. et al. Serial genomic inversions induce tissue-specific architectural stripes, gene misexpression and congenital malformations. *Nat. Cell Biol.* **21**, 305–310 (2019).
61. Despang, A. et al. Functional dissection of the *Sox9–Kcnj2* locus identifies nonessential and instructive roles of TAD architecture. *Nat. Genet.* **51**, 1263–1271 (2019).
62. Beagan, J. A. & Phillips-Cremins, J. E. On the existence and functionality of topologically associating domains. *Nat. Genet.* **52**, 8–16 (2020).
63. Alexander, J. M. et al. Live-cell imaging reveals enhancer-dependent *Sox2* transcription in the absence of enhancer proximity. *eLife* **8**, e41769 (2019).
This paper describes two-point live imaging in mouse embryonic stem cells with live readout of mRNA activity, and found no correlation of E-P proximity and promoter firing, in parallel work to Chen et al. (2018).
64. Huang, H. et al. CTCF mediates dosage- and sequence-context-dependent transcriptional insulation by forming local chromatin domains. *Nat. Genet.* **53**, 1064–1074 (2021).
65. Flavahan, W. A. et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110–114 (2016).
66. Phanstiel, D. H. et al. Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol. Cell* **67**, 1037–1048 (2017).
67. Bonev, B. et al. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572.e24 (2017).
68. Klein, D. C. & Hainer, S. J. Genomic methods in profiling DNA accessibility and factor localization. *Chromosome Res.* **28**, 69–85 (2020).
69. Krietenstein, N. et al. Ultrastructural details of mammalian chromosome architecture. *Mol. Cell* **78**, 554–565.e7 (2020).
70. Hsieh, T.-H. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell* **78**, 539–553.e8 (2020).
71. Hsieh, T.-H. S. et al. Enhancer–promoter interactions and transcription are maintained upon acute loss of CTCF, cohesin, WAPL, and YY1. Preprint at *bioRxiv* https://doi.org/10.1101/2021.07.14.452365 (2021).
72. Aljahani, A. et al. Analysis of sub-kilobase chromatin topology reveals nano-scale regulatory interactions with variable dependence on cohesin and CTCF. *Nat. Commun.* **13**, 2139 (2022).
73. Goel, V. Y., Huseyin, M. K. & Hansen, A. S. Region capture micro-C reveals coalescence of enhancers and promoters into nested microcompartments. Preprint at *bioRxiv* https://doi.org/10.1101/2022.07.12.499637 (2022).
74. Levo, M. et al. Transcriptional coupling of distant regulatory genes in living embryos. *Nature* **605**, 754–760 (2022).
75. Batut, P. J. et al. Genome organization controls transcriptional dynamics during development. *Science* **375**, 566–570 (2022).
76. Quinodoz, S. A. et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757.e24 (2018).
77. Beagrie, R. A. et al. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519–524 (2017).
78. Winick-Ng, W. et al. Cell-type specialization is encoded by specific chromatin topologies. *Nature* **599**, 684–691 (2021).
79. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
80. Barbieri, M. et al. Complexity of chromatin folding is captured by the strings and binders switch model. *Proc. Natl Acad. Sci. USA* **109**, 16173–16178 (2012).
This article introduced the 'strings and binders' model of chromatin folding, among the first models for TADs.
81. Jost, D., Carrivain, P., Cavalli, G. & Vaillant, C. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res.* **42**, 9553–9561 (2014).
82. Shi, G., Liu, L., Hyeon, C. & Thirumalai, D. Interphase human chromosome exhibits out of equilibrium glassy dynamics. *Nat. Commun.* **9**, 3161 (2018).
83. Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).
This article introduced the loop extrusion model to explain TADs, in parallel work with Sanborn et al. (2015).
84. Sanborn, A. L. et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl Acad. Sci. USA* **112**, E6456–E6465 (2015).
In parallel work to Fudenberg et al. (2016), this article introduced a tension globule and loop extrusion model to explain TADs and provided one of the first tests of the border function of single, directed CTCF sites.
85. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl Acad. Sci.* **115**, E6697–E6706 (2018).
86. Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harb. Perspect. Biol.* **2**, 38 (2010).
87. Chang, L.-H., Ghosh, S. & Noordermeer, D. TADs and their borders: free movement or building a wall? *J. Mol. Biol.* **432**, 643–652 (2019).
88. Szabo, Q. et al. TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci. Adv.* **4**, eaar8082 (2018).
89. Sikorska, N. & Sexton, T. Defining functionally relevant spatial chromatin domains: it is a TAD complicated. *J. Mol. Biol.* **432**, 653–664 (2020).
90. Nagano, T. et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* **547**, 61–67 (2017).
91. Flyamer, I. M. et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110–114 (2017).
92. Bonchuk, A. et al. Functional role of dimerization and CP190 interacting domains of CTCF protein in *Drosophila melanogaster*. *BMC Biol.* **13**, 63 (2015).
93. Chetverina, D., Aoki, T., Erokhin, M., Georgiev, P. & Schedl, P. Making connections: Insulators organize eukaryotic chromosomes into independent *cis*-regulatory networks. *Bioessays* **36**, 163–172 (2014).
94. Fujioka, M., Mistry, H., Schedl, P. & Jaynes, J. B. Determinants of chromosome architecture: insulator pairing in *cis* and in *trans*. *PLoS Genet.* **12**, e1005889 (2016).
95. Ghavi-Helm, Y. et al. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**, 96–100 (2014).
96. Beagan, J. A. et al. Three-dimensional genome restructuring across timescales of activity-induced neuronal gene expression. *Nat. Neurosci.* **23**, 707–717 (2020).
97. Ou, H. D. et al. ChromEMT: visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* **357**, eaag0025 (2017).
98. Sigal, Y. M., Zhou, R. & Zhuang, X. Visualizing and discovering cellular structures with super-resolution microscopy. *Science* **361**, 880–887 (2018).
99. Huang, B., Babcock, H. & Zhuang, X. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* **143**, 1047–1058 (2010).
100. Schermelleh, L., Heintzmann, R. & Leonhardt, H. A guide to super-resolution fluorescence microscopy. *J. Cell Biol.* **190**, 165–175 (2010).
101. Boettiger, A. & Murphy, S. Advances in chromatin imaging at kilobase-scale resolution. *Trends Genet.* **36**, 273–287 (2020).
102. Boettiger, A. N. et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
103. Szabo, Q. et al. Regulation of single-cell genome organization into TADs and chromatin nanodomains. *Nat. Genet.* **52**, 1151–1157 (2020).
104. Giorgetti, L. et al. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* **157**, 950–963 (2014).
105. Cattoni, D. I. et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat. Commun.* **8**, 1753 (2017).
106. Fudenberg, G. & Imakaev, M. FISH-ing for captured contacts: towards reconciling FISH and 3C. *Nat. Methods* **14**, 673–678 (2017).
This article provides an excellent comparison of FISH and Hi-C measurements, using both theory and experiment.
107. Wang, S. et al. Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016).
This article introduced chromosome tracing (1-Mb resolution), enabling validation of Hi-C compartments and analysis of compartments at the single-cell level.
108. Mateo, L. J. et al. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **568**, 49–54 (2019).
This paper introduced the ORCA chromosome tracing approach (up to 2-kb resolution), combined with RNA labelling to map cell-type specific changes in chromatin structure during embryonic development.
109. Cardozo Gizzi, A. M. et al. Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell* **74**, 212–222.e5 (2019).
This paper introduced the Hi-M chromosome tracing in whole embryos, combined with antibody detection of RNA.
110. Liu, M. et al. Multiplexed imaging of nucleome architectures in single cells of mammalian tissue. *Nat. Commun.* **11**, 2907 (2020).
This article introduced the MINA chromosome tracing approach in mouse embryonic tissue, combined with RNA and protein imaging.
111. Su, J.-H., Zheng, P., Kinrot, S. S., Bintu, B. & Zhuang, X. Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. *Cell* **182**, 1641–1659.e26 (2020).
In parallel work to Takei et al. (2021), this paper introduced the multiplexed barcoding DNA MERFISH method, locating over 1000 unique genomic loci throughout the genome in combination with genome-scale detection of nascent transcription and major protein markers of nuclear organization.
112. Takei, Y. et al. Integrated spatial genomics reveals global architecture of single nuclei. *Nature* **590**, 344–350 (2021).
In parallel work to Su et al. (2020), this article introduced the multiplexed barcoding SeqFISH+ method, locating 3,600 unique genomic loci, combined with labelling nascent RNA and protein.

113. Takei, Y. et al. Single-cell nuclear architecture across cell types in the mouse brain. *Science* **374**, 586–594 (2021).
114. Cardozo Gizzi, A. M. et al. Direct and simultaneous observation of transcription and chromosome architecture in single cells with Hi-M. *Nat. Protoc.* **15**, 840–876 (2020).
115. Liu, M. et al. Chromatin tracing and multiplexed imaging of nucleome architectures (MINA) and RNAs in single mammalian cells and tissue. *Nat. Protoc.* **16**, 2667–2697 (2021).
116. Mateo, L. J., Sinnott-Armstrong, N. & Boettiger, A. N. Tracing DNA paths and RNA profiles in cultured cells and tissues with ORCA. *Nat. Protoc.* **16**, 1647-1713 (2021).
117. Moffitt, J. R. et al. High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc. Natl Acad. Sci. USA* **113**, 11046–11051 (2016).
118. Rouhanifard, S. H. et al. ClampFISH detects individual nucleic acid molecules using click chemistry-based amplification. *Nat. Biotechnol.* **37**, 84–89 (2019).
119. Bintu, B. et al. Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362**, eaau1783 (2018).
This article reported improved resolution of chromosome tracing (30-kb), enabling validation of TADs and showing TADs to be statistical features emerging from averaging heterogeneous, TAD-like folds at the single-cell level.
120. Nir, G. et al. Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLoS Genet.* **14**, e1007872 (2018).
121. Hafner, A., Park, M., Berger, S. E., Nora, E. P. & Boettiger, A. N. Loop stacking organizes genome folding from TADs to chromosomes. Preprint at *bioRxiv* https://doi.org/10.1101/2022.07.13.499982 (2022).
122. Finn, E. H. et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* **176**, 1502–1515.e10 (2019).
123. Maeda, R. K. & Karch, F. The open for business model of the bithorax complex in *Drosophila*. *Chromosoma* **124**, 293–307 (2015).
124. Nicodemi, M. & Prisco, A. Thermodynamic pathways to genome spatial organization in the cell nucleus. *Biophys. J.* **96**, 2168–2177 (2009).
125. Mirny, L. A. The fractal globule as a model of chromatin architecture in the cell. *Chromosome Res.* **19**, 37–51 (2011).
126. Nicodemi, M. & Bianco, S. Chromosomes phase transition to function. *Biophys. J.* **119**, 724–725 (2020).
127. Mirny, L. A., Imakaev, M. & Abdennur, N. Two major mechanisms of chromosome organization. *Curr. Opin. Cell Biol.* **58**, 142–152 (2019).
128. Szabo, Q., Bantignies, F. & Cavalli, G. Principles of genome folding into topologically associating domains. *Sci. Adv.* **5**, eaaw1668 (2019).
129. Bickmore, W. A. V. The spatial organization of the human genome. *Annu. Rev. Genomics Hum. Genet.* **14**, 67–84 (2013).
130. Bickmore, W. A. & Van Steensel, B. Genome architecture: domain organization of interphase chromosomes. *Cell* **152**, 1270–1284 (2013).
131. Yildirim, A., Boninsegna, L., Zhan, Y. & Alber, F. Uncovering the principles of genome folding by 3D chromatin modeling. *Cold Spring Harb. Perspect. Biol.* **14**, a039693 (2022).
132. Conte, M. et al. Polymer physics indicates chromatin folding variability across single-cells results from state degeneracy in phase separation. *Nat. Commun.* **11**, 3289 (2020).
133. Espinola, S. M. et al. *Cis*-regulatory chromatin loops arise before TADs and gene activation, and are independent of cell fate during early *Drosophila* development. *Nat. Genet.* **53**, 477–486 (2021).
This paper reported that 'loop-dots' at the *docs* gene cluster form before TADs and are stable independently of transcription state.
134. Finn, E. H. & Misteli, T. Molecular basis and biological function of variability in spatial genome organization. *Science* **365**, eaaw9498 (2019).
135. Khanna, N., Zhang, Y., Lucas, J. S., Dudko, O. K. & Murre, C. Chromosome dynamics near the sol–gel phase transition dictate the timing of remote genomic interactions. *Nat. Commun.* **10**, 2771 (2019).
136. Chen, H. et al. Dynamic interplay between enhancer-promoter topology and gene activity. *Nat. Genet.* **50**, 1296–1303 (2018).
This paper describes two-point live imaging in *D. melanogaster* embryos with live readout of mRNA activity, and demonstrates strong correlation of E-P proximity and promoter firing for a synthetic transgene, within 4–8 minutes, in parallel work to Alexander et al. (2019).
137. Gabriele, M. et al. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science* **376**, 496–501 (2022).
This article describes live-cell imaging that shows that chromatin loops are highly dynamic and short-lived, in parallel work to Mach et al. (2022).
138. Mach, P. et al. Live-cell imaging and physical modeling reveal control of chromosome folding dynamics by cohesin and CTCF. Preprint at *bioRxiv* https://doi.org/10.1101/2022.03.03.482826 (2022).
This paper describes live-cell imaging that shows that chromatin loops are highly dynamic and short-lived, in parallel work to Gabriele et al. (2022).
139. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320.e24 (2017).
140. Abramo, K. et al. A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat. Cell Biol.* **21**, 1393–1402 (2019).
141. Zhang, H. et al. Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* **576**, 158–162 (2019).
142. Oudelaar, A. M. et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat. Genet.* **50**, 1744–1751 (2018).
143. Oudelaar, A. M. et al. A revised model for promoter competition based on multi-way chromatin interactions at the α-globin locus. *Nat. Commun.* **10**, 5412 (2019).
144. Allahyar, A. et al. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat. Genet.* **50**, 1151–1160 (2018).
145. Tavares-Cadete, F., Norouzi, D., Dekker, B., Liu, Y. & Dekker, J. Multi-contact 3C reveals that the human genome during interphase is largely not entangled. *Nat. Struct. Mol. Biol.* **27**, 1105–1114 (2020).
146. Cheutin, T. & Cavalli, G. The multiscale effects of polycomb mechanisms on 3D chromatin folding. *Crit. Rev. Biochem. Mol. Biol.* **54**, 399–417 (2019).
147. Zenk, F. et al. HP1 drives de novo 3D genome reorganization in early *Drosophila* embryos. *Nature* **593**, 289–293 (2021).
148. Weber, C. M. et al. mSWI/SNF promotes Polycomb repression both directly and through genome-wide redistribution. *Nat. Struct. Mol. Biol.* **28**, 501–511 (2021).
149. Kim, J. & Kingston, R. E. The CBX family of proteins in transcriptional repression and memory. *J. Biosci.* **45**, 16 (2020).
150. Narlikar, G. J. Phase-separation in chromatin organization. *J. Biosci.* **45**, 5 (2020).
151. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
152. Strom, A. R. & Brangwynne, C. P. The liquid nucleome–phase transitions in the nucleus at a glance. *J. Cell Sci.* **132**, jcs235093 (2019).
153. Finn, E. H. & Misteli, T. A genome disconnect. *Nat. Genet.* **51**, 1205–1206 (2019).
154. Xiao, J. Y., Hafner, A. & Boettiger, A. N. How subtle changes in 3D structure can create large changes in transcription. *eLife* **10**, e64320 (2021).
155. Zuin, J. et al. Nonlinear control of transcription through enhancer–promoter interactions. *Nature* **604**, 571–577 (2022).
156. Rajpurkar, A. R., Mateo, L. J., Murphy, S. E. & Boettiger, A. N. Deep learning connects DNA traces to transcription to reveal predictive features beyond enhancer–promoter contact. *Nat. Commun.* **12**, 3423 (2021).
157. Mir, M., Bickmore, W., Furlong, E. E. M. & Narlikar, G. Chromatin topology, condensates and gene regulation: shifting paradigms or just a phase? *Development* **146**, dev182766 (2019).
158. Payne, A. C. et al. In situ genome sequencing resolves DNA sequence and structure in intact biological samples. *Science* **371**, eaay3446 (2021).
159. Xie, L. et al. 3D ATAC-PALM: super-resolution imaging of the accessible genome. *Nat. Methods* **17**, 430–436 (2020).
160. Lu, T., Ang, C. E. & Zhuang, X. Spatially resolved epigenomic profiling of single cells in complex tissues. Preprint at *bioRxiv* https://doi.org/10.1101/2022.02.17.480825 (2022).
161. Nguyen, H. Q. et al. 3D mapping and accelerated super-resolution imaging of the human genome using in situ sequencing. *Nat. Methods* **17**, 822–832 (2020).
162. Jia, B. B., Jussila, A., Kern, C., Zhu, Q. & Ren, B. A spatial genome aligner for multiplexed DNA-FISH. Preprint at *bioRxiv* https://doi.org/10.1101/2022.03.25.485845 (2022).
163. Atkinson, T. J. & Halfon, M. S. Regulation of gene expression in the genomic context. *Comput. Struct. Biotechnol. J.* **9**, e201401001 (2014).
164. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.* **19**, 621–637 (2018).
165. Juven-Gershon, T. & Kadonaga, J. T. Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.* **339**, 225–229 (2010).
166. Halfon, M. S. Studying transcriptional enhancers: the founder fallacy, validation creep, and other biases. *Trends Genet.* **35**, 93–103 (2019).
167. Catarino, R. R., Neumayr, C. & Stark, A. Promoting transcription over long distances. *Nat. Genet.* **49**, 972–973 (2017).
168. Deng, W., Shi, X., Tjian, R., Lionnet, T. & Singer, R. H. CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc. Natl Acad. Sci. USA* **112**, 11870–11875 (2015).
169. Wang, Y. et al. Genome oligopaint via local denaturation fluorescence in situ hybridization. *Mol. Cell* https://doi.org/10.1016/j.molcel.2021.02.011 (2021).
170. Brown, J. M. et al. A tissue-specific self-interacting chromatin domain forms independently of enhancer–promoter interactions. *Nat. Commun.* **9**, 3849 (2018).
171. Beckwith, K. S. et al. Visualization of loop extrusion by DNA nanoscale tracing in single human cells. Preprint at *bioRxiv* https://doi.org/10.1101/2021.04.12.439407 (2021).
172. Karr, J. P., Ferrie, J. J., Tjian, R. & Darzacq, X. The transcription factor activity gradient (TAG) model: contemplating a contact-independent mechanism for enhancer–promoter communication. *Genes Dev.* **36**, 7–16 (2022).
173. Barinov, L., Ryabichko, S., Bialek, W. & Gregor, T. Transcription-dependent spatial organization of a gene locus. Preprint at *arXiv* https://doi.org/10.48550/arXiv.2012.15819 (2020).

Acknowledgements
The authors thank the members of the Boettiger laboratory for critical discussions and C. Walker for detailed feedback on the manuscript.

Author contributions
Both authors researched, discussed, wrote and edited the manuscript before submission.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Genetics thanks M. Nollmann and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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