Feature Review

CellPress

Unraveling the 3D genome: genomics tools for multiscale exploration

Viviana I. Risca and William J. Greenleaf

Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA

A decade of rapid method development has begun to yield exciting insights into the 3D architecture of the metazoan genome and the roles it may play in regulating transcription. Here we review core methods and new tools in the modern genomicist's toolbox at three length scales, ranging from single base pairs to megabase-scale chromosomal domains, and discuss the emerging picture of the 3D genome that these tools have revealed. Blind spots remain, especially at intermediate length scales spanning a few nucleosomes, but thanks in part to new technologies that permit targeted alteration of chromatin states and time-resolved studies, the next decade holds great promise for hypothesis-driven research into the mechanisms that drive genome architecture and transcriptional regulation.

The physical landscape of genome biology

More than a decade after the completion of a high-quality reference sequence [1,2], we have seen substantial strides in identifying elements of the genome that function in various cellular contexts. However, our understanding of the physical and spatial organization of the human genome at multiple scales remains stubbornly incomplete. The physical genome is an effectively 1D object housed in a nucleus 400 000 times shorter than its longest axis. This fundamental spatial constraint necessitates a complex hierarchical compaction of the 1D genome into 3D chromatin within the nuclear volume. The specifics of this 3D organization set the physical landscape of genome biology – from determining which regulatory elements are accessible to transcription factors (TFs) to regulating which distal enhancer elements make contact with which genes. By analogy to protein structure, the spatial hierarchy of genomic compaction can be roughly divided into three structured length scales: primary structure, comprising sequence elements, DNA-bound proteins, and nucleosomes; secondary structure, comprising interactions between nearby nucleosomes that shape local chromatin architecture; and tertiary structure, comprising long-range 3D features such as loops spanning a few hundreds of kilobases to chromosome domains spanning megabases (Figure 1) [3]. Thanks to the rapidly dropping cost of sequencing, novel genome-wide methods have been developed to begin to probe chromatin structure, especially at the primary and tertiary levels. An integrative, comprehensive, multiscale understanding of chromatin organization and a picture of how these layers of hierarchal organization bring about gene regulation is perhaps the next great challenge of genomics. In this review we briefly describe how the three levels of genome architecture impact transcriptional regulation and highlight some of the methodological developments that have provided these insights, focusing on methods that can probe 3D structure (Figure 1). Finally, we discuss the crucial blind spots in our understanding of the genomic landscape and outline the characteristics of future technologies that are still missing from the genomicist's toolbox.

Transcriptional regulation in three dimensions

Eukaryotic transcription is intimately tied to chromatin structure at multiple scales. The transcription machinery assembles on the core promoter, a primary-structure feature located around the transcription start site (TSS) of a gene that contains binding sites for components of the RNA polymerase pre-initiation complex [4]. Alone, the core promoter of most genes cannot activate reliable transcription because of the obstacle posed by chromatin [5]. Full transcriptional activation requires the action of TFs, DNA-binding proteins that also bind components of the RNA polymerase II transcription initiation complex and help to overcome this obstacle by recruiting chromatin-modifying and -remodeling complexes. The local secondary structure of chromatin, which may be altered by the action of these complexes, is thought to regulate both transcription and TF binding by controlling the accessibility of DNA [6,7]. Both gene expression profiles and genome-wide maps of open chromatin sites are highly cell-type specific [8].

Long-range tertiary interactions also play a role in regulating transcription. Although some TF-binding sites occur in or near core promoters of highly transcribed housekeeping genes, most TF-binding sites occur in enhancer regions. Enhancers are regulatory sequences that act in *cis* despite being as far as hundreds of kilobases away from their target genes [9,10]. The currently dominant, if not uncontroversial, model is that enhancer sequences form 3D contacts with promoters, integrating information from multiple TFs and tethering those TFs and their chromatin-remodeling binding partners to target promoters [11]. A natural question thus arises: what determines which enhancers contact which promoters? 3D genomics aims to answer this question by mapping the megabase-scale chromosomal interaction domains that are thought to

Corresponding author: Greenleaf, W.J. (wjg@stanford.edu).

Keywords: genome architecture; transcriptional regulation; chromatin structure; genomics methods.

0168-9525/

 http://dx.doi.org/10.1016/j.tig.2015.03.010

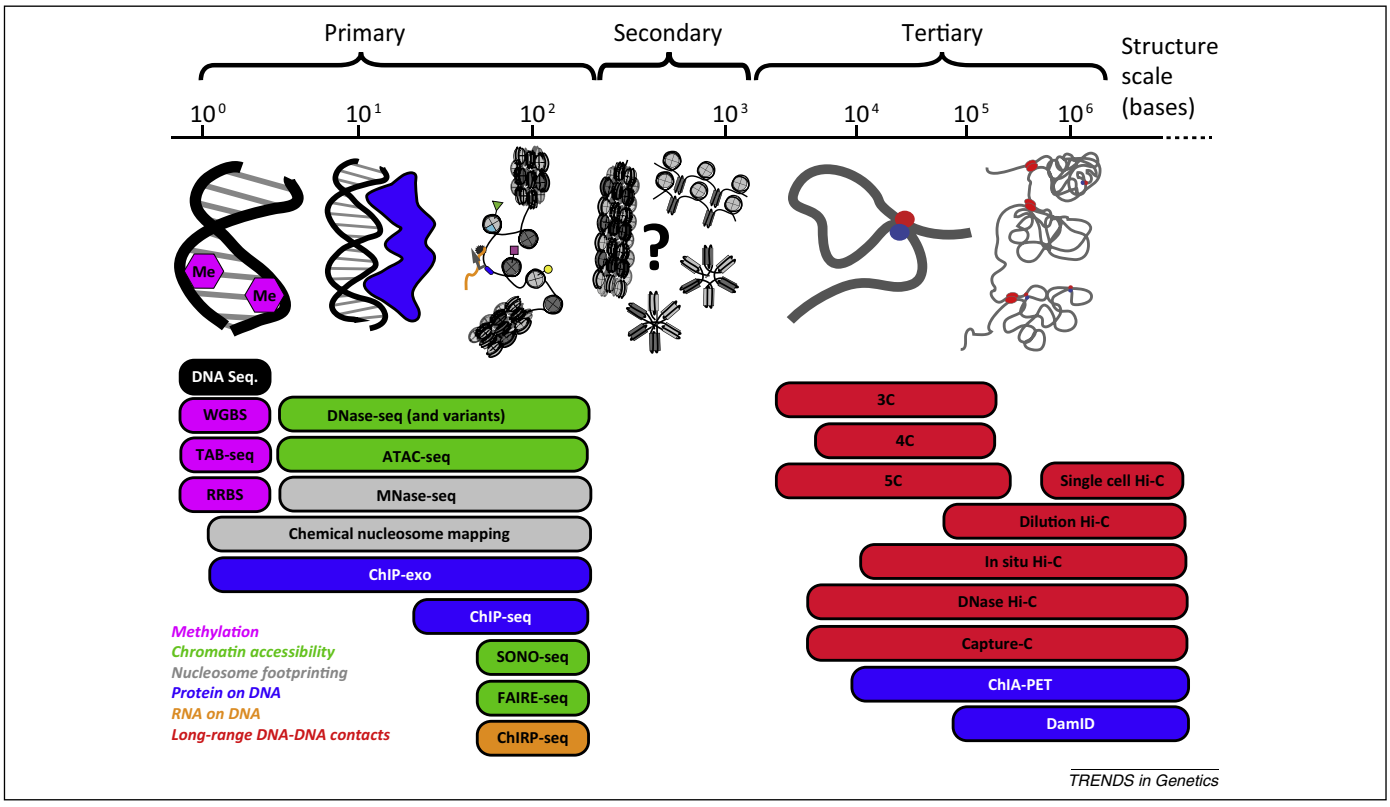

Figure 1. Overview of chromatin structure and assays at three scales. Chromatin structure is divided into three size scales, by analogy with protein structure. Primary structure encompasses DNA methylation (pink) and sequence features, DNA-bound factors (blue), nucleosome position and modifications (multicolored), and DNA accessibility. Secondary structure encompasses local structures formed by nucleosome–nucleosome interactions; although several models are shown here, the lack of methods that can probe this organizational scale of chromatin means that sequence-resolved *in vivo* architecture at this scale is not fully understood. Tertiary structure encompasses promoter–enhancer 3D contacts (spanning several kilobases to several hundreds of kilobases) and megabase-scale chromosome domains. Many methods exist for assaying the primary and tertiary scales of chromatin structure for both architecture and the identity of DNA-bound *trans* factors but no genomics methods directly assay secondary structure.

play a role in regulating enhancer–promoter looping [12]. The interaction domain map appears to correlate with genome-wide patterns of secondary and primary structure features like chromatin compaction and histone modifications [13], suggesting that the three length scales of chromatin architecture may feed back on each other, coordinating their regulation of transcription [14]. Thus, with the recognition over the past decade of the importance of the physical organization of DNA, a growing toolbox of methods has been developed to study transcriptional regulation at all three length scales. The resulting data sets are slowly generating a comprehensive, genome-wide picture of multiple aspects of this organization.

Primary structure

The primary structure of DNA comprises, first and foremost, DNA sequence. Analysis of the sequence can reveal TF-binding sites, methylation sites, and, in some cases, nucleosome positions, as well as identifying mutations between individuals. Primary structure also encompasses the epigenomic features that can be linearly mapped onto sequence, including histone modifications and variants and DNA accessibility.

At the smallest length scale of epigenomics, bisulfite sequencing [15,16] and its targeted variant reduced representation bisulfite sequencing (RRBS) [17] (Figure 2A and Table 1) have revealed that most of the human genome is stably methylated on CpG dinucleotides [18]; related assays have shown that hydroxymethylation may be important in brain tissue [19,20]. Regions rich in hypomethylated CpGs coincide with some promoters and have been proposed to act as seeding regions for open chromatin, allowing early-acting pioneer TFs to bind DNA and locally open up the chromatin for other factors [5,21].

At the length scale of a few to a few hundred base pairs, methods can be divided into those that probe chromatin accessibility and those that map the binding of a factor of interest on a linear DNA sequence (Figure 2D). In the first category, formaldehyde-assisted isolation of regulatory elements (FAIRE)-seq sequences accessible DNA by depleting nucleosome-associated DNA [22]. DNase-seq [23] uses the endonuclease DNase I to preferentially nick DNA in open chromatin then sequences the resulting fragments [6,24,25]. A new method, assay for transposase-accessible chromatin (ATAC)-seq, uses the Tn5 transposase to fragment DNA at open chromatin regions and efficiently produces sequencing libraries in a single step from low cell numbers [26]. Regardless of which method is used, open chromatin analysis is a highly effective way to study transcriptional regulation genome wide, as nuclease-sensitive sites have long been known to indicate active regulatory regions such as promoters and enhancers [27,28]. Indeed, one approach for assigning enhancers to their target promoters has been to look for correlations in open chromatin at pairs of enhancers and promoters across cell types [8].

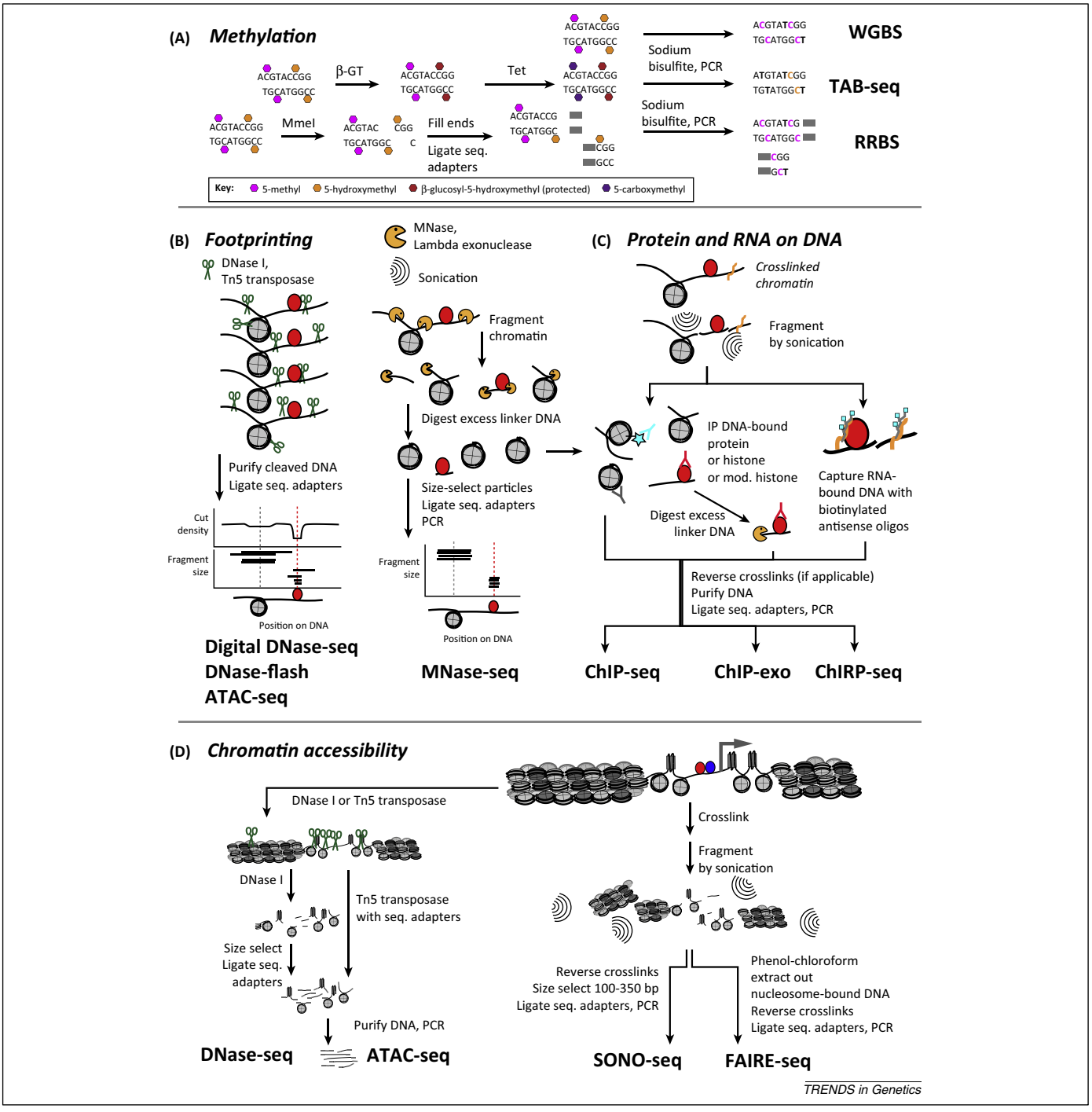

Figure 2. Methods that assay the primary structure of chromatin. (**A**) The most widely used assays for cytosine methylation with base pair accuracy rely on sodium bisulfite, which converts unprotected cytosines to uracils. In a sequencing reaction, methylated and hydroxymethylated sites are read as cytosines while all other cytosines are read as thymines. Tet-assisted bisulfite sequencing uses two additional enzymatic steps to selectively protect hydroxymethylcytosines from bisulfite conversion and reduced-representation bisulfite sequencing uses a methylation-sensitive restriction enzyme to cleave near methylated CpGs, ensuring that they are read during sequencing. (**B**) DNA footprinting can be done with either endonucleases or transposases, which cleave unprotected DNA, or with MNase, which has hybrid endonuclease and exonuclease activity and 'nibbles' free DNA until it reaches an obstacle such as a transcription factor (TF) or nucleosome. DNA fragments are purified, size selected, and sequenced. DNase-seq focuses on TF footprints from short fragments whereas DNase I-released fragment-length analysis of hypersensitivity (DNase-FLASH) isolates multiple size classes to assay both nucleosome and TF footprints. (**C**) Fragmented chromatin can be assayed to identify the sequence of DNA bound to protein (ChIP) or RNA [chromatin isolation by RNA purification (ChIRP)]. The fragmentation methods vary, from sonication of crosslinked chromatin to micrococcal nuclease (MNase) digestion of uncrosslinked chromatin (called native ChIP). Target proteins are pulled down with antibodies whereas RNAs of interest are pulled down with biotinylated antisense oligos. ChIP-exo employs lambda exonuclease to digest away excess free DNA that is pulled down, increasing the resolution of the method. (**D**) DNA accessibility, a useful indicator of active regulatory regions, can be assayed by DNase I cleavage [DNase-seq/DNase I hypersensitive site (DHS)-seq], Tn5 transposase insertion of sequencing adapters [assay for transposase-accessible chromatin (ATAC)-seq], fragmentation by sonication (Sono-seq), or the depletion of protein-bound fragments [formaldehyde-assisted isolation of regulatory elements (FAIRE)-seq]. A 30-nm chromatin fiber cartoon is used to depict the contrast between open and closed chromatin in this panel.

Table 1. Selected methods for primary chromatin structure

Method	Assay for	Description	Resolution (bp)	Required input	Notes
Whole-genome bisulfite sequencing (WGBS) [15,16]	Cytosine methylation and hydroxylmethylation	Sodium bisulfite converts unprotected cytosines to uracil. Sample is sequenced with and without bisulfite treatment.	1	20 ng genomic DNA (gDNA)	This assay cannot distinguish methylation from hydroxymethylation. Several updates have improved this protocol for small numbers of cells [15,16].
Tet-assisted bisulfite sequencing [20] and oxidative bisulfite sequencing [19]	Cytosine hydroxylmethylation	Hydroxymethylcytosine is glucosylated by beta-glucosyltransferase and methylcytosine is converted to carboxymethylcytosine by Tet. On treating these libraries with sodium bisulfite and sequencing, only the hydroxymethylcytosine reads as C [20]. Alternatively, hydroxymethylcytosine is oxidized to formylcytosine, which is converted to uracil by bisulfite treatment [19].	1	3 μg gDNA	Recent studies have addressed the role of hydroxymethylation in brain tissue, where it is particularly enriched, and the function of endogenous Tet proteins in regulating chromatin [20].
RRBS [17]	Cytosine methylation and hydroxylmethylation	A methylation-insensitive RE that cuts at CpG dinucleotides is used to generate fragments that contain at least one CpG site.	1	100–300 ng gDNA	This protocol is more cost-efficient than WGBS because it enriches for reads containing potential methylation sites that are of interest.
FAIRE-seq [22]	Open chromatin	Nucleosome-bound DNA is crosslinked and removed by phenol–chloroform extraction and the remaining nucleosome-free DNA is analyzed by microarray or sequencing.	~200	10^6–10^7	This is a simpler assay for open chromatin than DNase-seq, although its resolution is somewhat lower.
Sonication of crosslinked chromatin (Sono)-seq [154]	Open chromatin	Crosslinked chromatin is sonicated as for a ChIP experiment (see below) but no immunoprecipitation is performed. This signal is compared with sonicated naked DNA.	100–350	$\sim10^8$	This method arose from an observation that no-immunoprecipitation input libraries in ChIP experiments had peaks in sonication breaks near promoters and other open chromatin sites.
DNase-seq or DNase I hypersensitive site (DHS)-seq [6,8,23–25,29–31,155,156]	Open chromatin and TF footprints	DNA in isolated nuclei is digested with DNase I at a concentration that must be optimized for each experiment. A library is prepared from the digested fragments by ligation of adapters and cleavage of ~20-bp sequence tags followed by size selection of a unique library molecule size or by biochemical fractionation of fragments followed by ligation of sequencing adapters.	~1	10^7–10^8	In one experiment, this method revealed the genome-wide landscape of open chromatin, which correlates well with all TF-bound sites. Technically difficult, involving optimization of DNase I concentration. The sequence bias of DNase I can give rise to artifactual apparent footprints on DNA [122].
DNase I-released fragment-length analysis of hypersensitivity (DNase-FLASH) [33]	Nucleosome and TF footprints	This variation on DNase-seq systematically analyzes a broad range of fragment sizes to capture both nucleosome footprints and TF footprints.	1	10^7–10^8	This method is particularly useful for mapping the structure of TSSs and regulatory regions because it simultaneously probes the positions of nucleosomes and TFs.
ATAC-seq [26]	Open chromatin, nucleosome positions, TF footprints	Tn5 transposase is used to transpose sequencing adapter oligos into the gDNA of permeabilized, unfixed cells. The resulting library is then purified and sequenced.	~1	10^2–10^4	This simple, fast protocol generates very similar data to DNase-seq with lower input requirements. Tn5 transposase also exhibits sequence bias, which must be accounted for in footprinting experiments [122].
MNase-seq [36,37,157]	Nucleosome footprints (and, in a variation, smaller DNA-bound proteins)	MNase is used to digest chromatin in nuclei from cells lysed by cryogenic grinding or detergent. The enzyme has endonuclease activity that cuts linker DNA between nucleosomes and exonuclease activity that digests any linker not protected by the core nucleosome. Nucleosome-sized fragments are	~1–10	10^7	MNase digestion is the method of choice for global fragmentation of chromatin into nucleosome core particles. Digestion conditions used in MNase-seq experiments do not select for cleavage in open chromatin regions only, generating data genome wide and thus requiring large numbers of reads to obtain sufficient depth. MNase is

Table 1 (*Continued*)

Method	Assay for	Description	Resolution (bp)	Required input	Notes
		size selected and prepared for sequencing.			often used to fragment nucleosomes before ChIP (see below).
Chemical mapping of nucleosomes [124,125]	High-resolution nucleosome position mapping	This method requires histone mutants that contain a single cysteine in the core nucleosome near the nucleosome dyad. A thiol-reactive copper chelator is chemically linked to the unique pair of thiols on nucleosomes in fixed and permeabilized cells. Copper is added and peroxide is used to oxidize the copper, locally generating free radicals that cleave the DNA at two stereotyped positions around the nucleosome dyad. The resulting DNA fragments are then purified and sequenced.	~1	Budding and fission yeast only	The free radical-mediated cleavage of DNA eliminates enzyme-based bias, although sequence bias may still be introduced by downstream library preparation steps. This method has currently been implemented in *Saccharomyces cerevisiae* [124] and *Schizosaccharomyces pombe* [125].
ChIP-seq [40]	Mapping of DNA-bound proteins (including nucleosomes)	Chromatin is fixed in cells then fragmented by sonication or MNase digestion before enrichment for the protein epitope of interest using a specific antibody. Crosslinks are reversed using proteinase K and heat and the DNA is then prepared for analysis by sequencing, array hybridization, or PCR.	~100	$\sim 10^7$	The resolution of this method is limited by fragment size, which is determined by sonication in experiments targeted at non-nucleosome proteins, and falls in the 100–200-bp range.
Native ChIP-seq [41,42,157–159]	Mapping of DNA-bound proteins under native conditions (mainly nucleosomes)	Chromatin is digested with MNase in permeabilized but unfixed cells and then solubilized by passing through fine-gauge needles before preparation of a sequencing library from the fragments. Mono- and dinucleosome fragments are isolated by density gradient centrifugation of nucleosome particles or by size selection of DNA after adapter ligation.	~10	$\sim 10^3$–10^7	This method relies on the stability of the nucleosome–DNA interaction to maintain the association of the bound DNA with the protein target of the antibody. One recent variation of this method uses low salt to stabilize the TF–DNA interaction for native TF ChIP [42] while another has optimized the protocol for input as low as 1000 cells [159].
ChIP-exo [160]	High-resolution footprinting of nucleosomes with particular modifications and DNA-bound proteins	Crosslinked and fragmented DNA with bound protein is digested with lambda exonuclease to remove excess unprotected DNA before reversal of crosslinks, library preparation, and single-end sequencing.	~1	10^6	The development of this method dramatically increased the resolution of the long-standing crosslinked ChIP protocol.
ChIRP-seq [53,54]	Mapping of genome interactions with a target lncRNA	Biotinylated antisense oligos are designed to tile the ncRNA of interest. These oligos are then hybridized to chromatin that has been gluteraldehyde crosslinked in intact cells then fragmented by sonication. Oligo-bound RNA and its associated DNA are isolated with streptavidin beads and the DNA is then purified and sequenced.	~100	$\sim 10^7$	This method permits mapping of RNA–DNA interactions far from the transcribed locus of the RNA of interest.
STARR-seq [51,52]	Functional genome-wide assay for enhancer activity	Fragments spanning the genome are cloned into a library of GFP reporter constructs between the polyadenylation site and the transcription termination site.	N/A	*Drosophila* only	STARR-seq reveals the potential of a fragment of genomic sequence to act as an enhancer for a promoter of interest. The results therefore depend on the promoter sequence used in the assay.

In areas of accessible chromatin, DNase, Tn5 transposase, and micrococcal nuclease (MNase) can also be used for footprinting assays, in which short fragments of DNA protected against nuclease cleavage by bound nucleosomes or TFs are purified and sequenced (Figure 2B) [26,29–33]. All of these enzymes can be used to map both nucleosomes and TFs, but because MNase digests all linker DNA while leaving nucleosome-bound DNA intact, this approach works well for genome-wide nucleosome footprinting [34–37], whereas DNase I is primarily used for TF footprinting because it can nick the minor groove of nucleosome-bound DNA [38]. These data must be interpreted with caution due to sequence bias in the digestion preferences of enzymes (see 'Outlook for 3D genomics' below).

In the second category, ChIP- and chromatin isolation by RNA purification (ChIRP)-seq have provided genome-wide maps of DNA-bound proteins and DNA-associated RNAs, respectively (Figure 2C and Table 1) [39,40]. ChIP-seq is usually performed using formaldehyde crosslinking to stabilize the association of proteins with DNA, but native ChIP using MNase to fragment chromatin is also possible for nucleosomes [41] or for TFs in low salt [42]. The quality of ChIP-seq experiments hinges on the specificity and sensitivity of the antibody. Large-scale systematic efforts to perform ChIP with carefully validated antibodies in multiple cell lines have generated an invaluable resource for the genomics community [43,44]. Histone modification ChIP data are useful both for identification of regulatory elements [45] and more generally as an input to computational strategies for determining chromatin states [46,47]. TF ChIP data sets have been fed into TF-binding motif databases [48–50] and combined with open chromatin data to create genome-wide maps of presumptive TF binding in cell lines where individual TF ChIP experiments were not necessarily performed [8,30]. Self-transcribing active regulatory region (STARR)-seq, a parallelized fluorescent reporter assay, does not assay primary structure *per se* but has been applied to the *Drosophila* genome to search genome wide for sequences that can drive transcription [51,52] and provides a complementary approach to the same problem (Table 1).

ChIRP-seq and a similar method called capture hybridization analysis of RNA targets (CHART)-seq use short biotinylated oligonucleotide probes that hybridize to an RNA of interest instead of an antibody [53,54]. ChIRP-seq has already been used to show the genome-wide interactomes of long noncoding RNAs (lncRNAs) involved in Polycomb repressive complex (PRC) 2-mediated silencing, telomere function, and dosage compensation [53,54].

Secondary structure

In mammalian genomes, where most nucleosomes are not well positioned [36,37], precise nucleosome positioning cannot account for regulation of the accessibility of TF-binding sites, so chromatin-compaction models have been invoked for the regulation of transcription [5]. We define secondary structure as the local nucleosome–nucleosome interactions that integrate many primary-structure features such as histone modifications and DNA-bound proteins to regulate the accessibility of DNA, the compaction of chromatin, and the physical properties of the chromatin fiber.

The proposed 30-nm fiber model of secondary structure [55–58], built from consecutive nucleosomes that collapse into a helix, is analogous to a protein alpha helix. The proposed chromatin melt model [59] suggests either zigzags of interdigitated non-consecutive nucleosomes, analogous to the beta sheet, or a completely unstructured chromatin fiber, analogous to the collapsed random coil. A fourth model of secondary structure is beads-on-a-string uncondensed chromatin, analogous to an unstructured peptide. The first three models are proposed for compact arrangements that preclude most TF binding and transcription activity, whereas the fourth is believed to be the open structure that facilitates these processes [5].

Currently, no sequencing-based methods can probe the secondary structure of unperturbed chromatin and report anything beyond a description of 'open' or 'closed' [60]. DNA-accessibility assays like DNase-seq and ATAC-seq report on an important effect of secondary structure but do not reveal any details about the 3D architecture of the compacted state *per se* or the mechanisms that maintain it. In the absence of higher-throughput and higher-resolution methods, the field has patched together the current understanding of chromatin secondary structure from four main approaches.

First, *in vitro* experiments with extracted or reconstituted chromatin have been used to study the modulation and regulation of chromatin secondary structure by cations [55,57,61], architectural proteins or complexes including linker histone H1 [62,63], heterochromatin protein 1 (HP1) [64], methyl-CpG-binding protein 2 (MeCP2) [65], and PRC1 [66], chromatin remodelers [67], nucleosome-positioning and linker histones [68], histone variants [69], and histone modifications, especially acetylation [70] under dilute conditions. Some functional experiments with reconstituted chromatin have also demonstrated how deacetylation and linker histone-dependent chromatin compaction blocks transcription [71]. Most of these *in vitro* experiments are consistent with a 30-nm fiber model of chromatin [57].

Second, many of the histone modifications and architectural proteins studied *in vitro* have been mapped onto the genome's sequence using ChIP and DNA adenine methyltransferase identification (DamID) (described below) [72], as have nucleosome positions, which determine the local internucleosome linker length [68]. Third, modeling based on known features of DNA elasticity and nucleosome structure have delineated the space of possible chromatin architectures [73] and shown that there is no unique minimum-energy solution, emphasizing the importance of regulatory factors and electrostatic interactions [58].

Fourth, electron microscopy and X-ray scattering methods have been used to probe gently fixed or frozen nuclei and have shown evidence for 30 nm fibers, but only in metabolically inactive cells [74–76]. Fluorescence *in situ* hybridization (FISH)-based measurements have shown compaction of chromatin between two loci, but the architecture of the compacted state is unknown [77]. A lack of evidence for 30 nm fibers in the nuclei of normal somatic cells suggests that the polymer melt model may be more

appropriate [3,59,75,78,79]. Resolution of this debate is likely to require an understanding of fiber polymorphism and regulation [80–82]. Discrepancies between *in vitro* and *in vivo* measurements of chromatin secondary structure are likely to arise from both the high density of nucleosomes and chromatin-binding and -remodeling factors found in the living nucleus. As new genomics methods are developed and the resolution of existing methods is increased, global maps of chromatin secondary structure should provide a framework to facilitate integration of these *in vivo* and *in vitro* data sets.

Tertiary structure

Tertiary structure in this context comprises long-range 3D chromatin contacts and organization. Mapping the tertiary structure of the genome (Figure 1) is a promising approach to understanding the physical wiring diagram of the genome, linking enhancers to the genes they regulate. Many of the methods for probing tertiary structure are based on proximity ligation, a technology first developed as part of chromosome conformation capture (3C) (Figure 3 and Table 2) to map the global 3D structure of a yeast chromosome [12,83,84]. The power of 3C to provide insights into transcriptional regulation was demonstrated with an investigation of looping between DNase I hypersensitive sites and their target gene at the mouse β-globin locus, which inspired the active chromatin hub model of transcriptional activation [85]. This hub is a pre-existing cluster of enhancers and the activation of transcription of target genes during differentiation is associated with the genes looping into the hub [86]. 3C is best used for exploring a single group of enhancers, such as the globin locus, or for validating the results of higher-throughput methods.

4C, chromosome conformation capture carbon copy (5C), and Capture-C methods (Table 2 and Figure 3B) are variants of 3C that allow 3D looping interactions to be assayed in parallel at multiple sets of loci [87–90]. They have proven useful in either mapping a region of interest, such as an extended locus with multiple promoters and enhancers that have developmentally regulated interac-

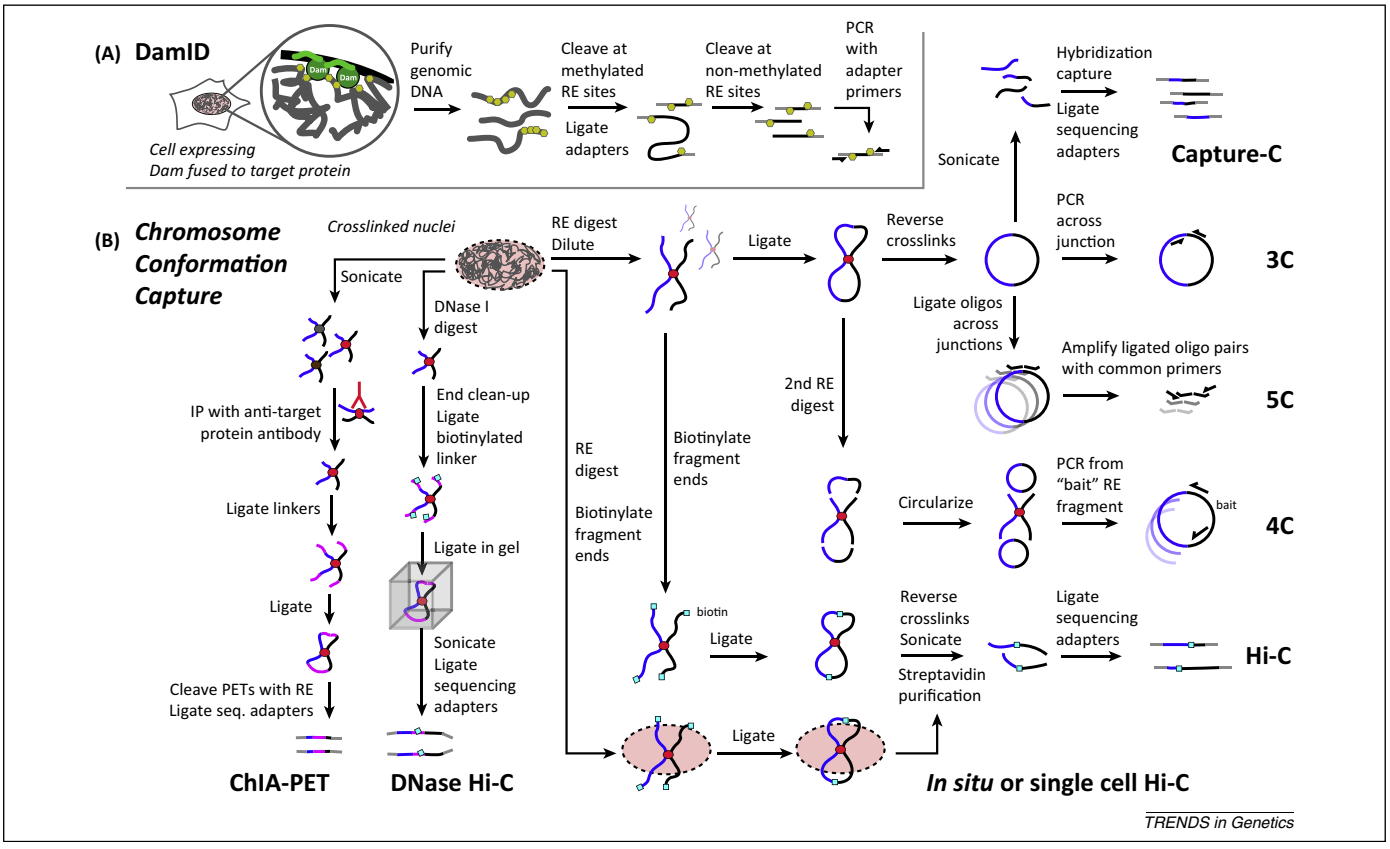

Figure 3. Methods that assay the tertiary structure of chromatin. (**A**) DNA adenine methyltransferase identification (DamID) relies on the expression of a fusion between the bacterial adenine methylase Dam and the protein of interest in cells. Dam methylates (yellow) TCGA motifs on any genomic DNA it contacts and genomic DNA is enzymatically cleaved at methylated restriction sites; the sites are then ligated to adapter oligos (gray) that are complementary to PCR primers. Another restriction digestion cleaves non-methylated sites. The amplification step selects for densely methylated areas, in which methylated restriction sites are adjacent. (**B**) The chromosome conformation capture (3C) family of methods is based on a restriction digest of crosslinked chromatin followed by proximity ligation of the fragment ends in dilute solution. Proximity ligation creates junctions between fragments from genomic regions that interact in 3D. Combining this approach with different strategies for fragmentation, enrichment, and read out of the ligation junctions yields a diverse set of methods. In 3C, PCR with primers on either side of the ligation junction of interest assay a single site. In 4C, the library is cleaved with a second restriction enzyme (RE) and recircularized. PCR primers oriented outward from the test (bait) locus are then used to amplify the unknown regions with which the bait interacts. In chromosome conformation capture carbon copy (5C), oligos are designed to bridge ligation junctions and are then hybridized to and ligated across the junctions to create 'carbon copies' of the junctions that can be PCR-amplified with common primers. Capture-C involves fragmentation of a 3C library then oligo hybridization capture to enrich for all sequence fragments (junctions and non-junctions) from a subset of the genome. Hi-C is a very high-throughput version of 3C that assays all interactions between all genomic loci. Fragment ends are biotinylated (light blue squares) before proximity ligation to facilitate their enrichment for sequencing library construction. Refinements of Hi-C include proximity ligation within crosslinked nuclei rather than in dilute solution and the use of DNase I fragmentation to replace restriction digest. Because DNase I generates various end types, the ends must be blunted and ligated to biotinylated linkers (pink with light blue square) to replace the biotinylation of sticky ends from RE digest. Chromatin interaction analysis with paired end tag (ChIA-PET) assays DNA–DNA contacts involving a particular protein by incorporating a ChIP step in the workflow before proximity ligation.

Table 2. Selected methods for assaying tertiary chromatin structure

Method	Assay type	Protocol description	Resolution	Input	Read out	Notes and example applications
3C [83]	One locus to one locus	Formaldehyde crosslinking, restriction digest, ligation in dilute solution, crosslink reversal	Depends on distance from loci of interest to nearest restriction site; as high as 1 kb	~10^8 yeast nuclei or ~10^7 mammalian cells	PCR	Sequence of bait locus and interacting locus must be known. Analysis is straightforward. • Early work on whole-chromosome 3D conformation in yeast [83]. • Comparing looping of enhancer regions to a promoter of interest between cell types [85] and through stages of differentiation [86].
				~10^4–10^7 cells	qPCR [161]	This is a more precise and sensitive approach than quantifying the amount of yield from PCR. • Measuring quantitative changes in interaction frequency between enhancers through differentiation, after genetic or knock down disruptions of the looping machinery [162], or during induction of ectopic looping [114]. • High-resolution validation of interactions identified by Hi-C [94].
4C [87,88,163]	One 'bait'/ 'viewpoint' locus to the genome	3C library is cut with a second RE and circularized, then PCR amplified with primers facing out from bait locus	Up to 2 kb sliding window (4C-seq); data are usually analyzed at multiple window sizes ('domainogram')	~5×10^5 cells	Microarray or seq. 1–2 M reads	Sequence of bait locus must be known. Provides the genomic interactome viewpoint from the bait locus. • Baits can be multiplexed for multiple viewpoints in one experiment [164].
5C [89]	Many loci to many loci	3C with ligation-mediated amplification (LMA) of a 'carbon copy' library of oligos designed across restriction fragment junctions of interest	Up to ~1 kb	~10^7 cells	Microarray [89] or seq. ~10 M reads per sample [165]	Loci to be assayed limited by complexity of LMA oligo pool. • High-resolution TAD mapping over a region of interest (4.5 Mb) [93]. • Systematic mapping of long-range looping interactions between TSSs and putative regulatory regions identified by ENCODE [165].
Capture-C [90]	Multiplexed one to one in regions of interest	3C library is sonicated and selected loci are captured with oligo hybridization	2 kb sliding window in selected regions of interest	~10^7 cells	Seq. 200 M reads per sample	Oligo capture targeted ~450 promoters. Most looping interactions occur within ~300 kb of the target promoter and within topological domains [90]. Can be used for large-scale analysis of SNPs on looping interactions.
Dilution Hi-C [91]	Whole genome to whole genome	3C restriction fragment ends are biotinylated before proximity ligation; library is sheared and biotinylated junctions are pulled down and sequenced	0.1–1 Mb [91], 40 kb [92], 5–10 kb [94]	~10^7 cells	Seq. 8.4M [91], 3.4 B [94] reads per sample	This is the original Hi-C method. It maps megabase-scale topological domains in mammalian genomes. Limited by sequencing depth and frequency/distribution of restriction sites in the genome. Biases are due to chromatin openness, sequence composition, and fragment length. • Mapped TADs in mouse and human genomes [92] whose borders are conserved and occupied by CTCF-binding sites and housekeeping genes. • Showed that genome-wide, cell type-specific looping interactions exist before induction of transcription and potentiate transcription after induction via tumor necrosis factor alpha (TNF-α) signaling [94]. • Characterized the distances between putative enhancers and promoters and the numbers of looping partners observed for each [94].
DNase Hi-C [95]	Whole genome to whole genome	Crosslinked DNA bound to beads is fragmented by DNase I instead of REs, then a biotinylated adapter is ligated before proximity ligation of	1–50 kb or 1 Mb	2–5 × 10^6 cells	Seq. ~100–800 M reads per sample	Resolution does not depend on locations of RE sites, limited only by sequencing depth, and thus can obtain better coverage than RE Hi-C with deep sequencing. High-resolution data was achieved using targeted amplification to focus on large intergenic ncRNA (lincRNA) promoters [95].

Table 2 (*Continued*)

Method	Assay type	Protocol description	Resolution	Input	Read out	Notes and example applications
		molecules embedded in agarose.				
In situ Hi-C [13]	Whole genome to whole genome	The early steps of Hi-C, including restriction digest and proximity ligation, are performed in permeabilized but otherwise intact nuclei, without dissociating chromatin.	1–5 kb	~2–5 × 10^6 cells	Seq. 112 M to 3.6 B reads per biological replicate (primary experiment)	Higher resolution and better signal to noise than dilution Hi-C. These relatively recent experiments also include much higher sequencing depth than most previous Hi-C studies. • Showed that many loops are mediated by CTCF bound to pairs of motifs in a convergent orientation [13]. • Classified genome domains defined by stable loops into chromatin states [13]. • Haplotype-specific mapping of imprinted loci and inactivated X chromosome conformation [13] using SNPs in Hi-C reads.
Single cell Hi-C [133]	Whole genome to whole genome	Same as *in situ* Hi-C, but single nuclei are first isolated and processed independently to add unique barcodes before amplification and sequencing	~10 Mb (single cells) to 1 Mb (pooled cells)	1–60 cells	Seq. ~10 K read pairs per cell	The first ligation-based method that can provide single-cell resolution for looping interactions (which is otherwise obtainable by FISH measurements). • Showed that TADs are stable and reproducible but inter-TAD and interchromosome contacts are stochastic [133].
ChIA-PET [103]	Whole genome to whole genome mediated by protein of interest	Chromatin is crosslinked and sheared by sonication and enriched by ChIP against a protein or histone modification of interest; biotinylated half-linkers are ligated to ends of DNA fragments, followed by proximity ligation of fragment ends; a RE is used to cleave ~20 bp on either side of the linker, then these paired-end tags are selected and sequenced	Depends on read depth and the size of the genome region bound by the protein of interest; ultimately limited by the size of the sheared chromatin fragments; ~3-kb fragments counted as self-ligation [103]	~10^8 cells	Seq. ~20–30 M read pairs per sample [103], 100–200 M read pairs per sample [106]	Restricts looping analysis to loops mediated by a protein of interest such as CTCF. • Used to study chromatin modification patterns in CTCF-mediated loops [104], loops induced on estrogen signaling and mediated by estrogen receptor alpha [103], loops involving RNA polymerase [105], and a combination of proteins that regulate genome architecture [106].
DamID [72,166]	Protein to whole genome	Dam is fused to a protein of interest and expressed in the cell. It methylates adenines at CATG sites in proximity to the protein of interest, which are then detected by methylation-specific PCR	~1 kb	~10^6 mammalian cells (~2.5 μg gDNA) [167]	Methylation-specific PCR [72] and microarray hybridization [167]	Can provide similar information to ChIP but the interaction occurs in live cells over a long time window. • Used to map the interactions of chromosomes with nuclear bodies, a major one being the lamina; used to define LADs [108]. • Also used to map DNA–protein interactions (e.g., Polycomb group proteins [168]). • Not suitable for mapping histone modifications [167].

tions, or surveying hundreds of promoter regions genome wide. In many of these approaches, the design of PCR primers or capture oligos can also select particular SNPs, probing a single allele [90].

The highest-throughput version of 3C is Hi-C, which has been used to systematically map pairwise interaction frequencies genome wide (Table 2 and Figure 3B). Its resolution is limited by the size of the restriction fragments used and by sequencing depth. Initial results were binned to a resolution of approximately 1 Mb to overcome the challenge of low genome coverage of sequenced ligation junctions [91]. At 1-Mb resolution, the human genome appears to partition into multi-megabase domains that belong to either of two effective nuclear compartments, termed A and B, that roughly correspond to active and inactive chromatin [91]. With deeper sequencing, it was possible to resolve topologically associating domains (TADs), regions spanning hundreds of kilobases to approximately 1 Mb defined by the higher likelihood that two loci lying in the same TAD will interact with each other rather than with loci in a different TAD [92,93]. The resolution of Hi-C has been pushed further still to reveal sub-TAD structure genome wide by employing very high sequencing depth combined with careful filtering of the data [94], by replacing restriction enzyme (RE) digestion with DNase I fragmentation to remove the restrictions imposed by restriction site frequency in the genome [95], and by using *in situ* proximity ligation, which helps to reduce nonspecific ligations in combination with extremely deep sequencing [13].

The 3C family of methods has rapidly advanced our understanding of 3D genomic architecture. Interestingly, a significant portion of TAD borders are conserved between species [92] and between cell types during differentiation [93] and their boundaries are not significantly altered after stimulation of major cell signaling pathways [94], suggesting that TADs may constitute a basal architecture of mammalian genomes on which finer-grained regulatory interactions are overlaid. TAD borders contain binding sites for the ubiquitous CTCF [13,92,93], which has been associated with an insulator function that blocks communication between enhancers and promoters when placed between them and prevents the spread of heterochromatin in reporter systems [96–98]. Expression profiles of genes in the same TAD are correlated during differentiation [93] and CTCF-bound TADs can be roughly clustered by the sets of genomic loci with which they interact [13]. TADs are also concordant with replication-timing domains [99], further supporting their role as functional modules of the genome. High-resolution 5C experiments that focused on sub-TAD structure using existing ChIP data observed that long-range 3D interactions are hierarchical [100]. CTCF and cohesin-mediated long-range interactions constitute the TADs that do not change appreciably as cells differentiate, whereas developmentally regulated, shorter-range contacts within TADs are associated with cohesin and the Mediator complex (a component of the transcription initiation machinery [101]). Interestingly, high-resolution Hi-C has shown that enhancer–promoter contacts are cell-type specific but, at least in some cases, can be detected before transcriptional activation by a cell signaling cascade is induced, suggesting that, in addition to constitutive TADs, which are shared across cell types, an additional layer of cell type-specific 3D interactions exists to potentiate the transcriptional activation of certain genes in response to exogenous signals [94,102].

The 3C family of methods focus only on DNA–DNA contacts and although they can be used to study the protein factors that regulate the tertiary structure of the genome in combination with ChIP, more direct methods for understanding the role of particular proteins in shaping the tertiary structure of chromatin have also been developed. Chromatin interaction analysis with paired end tag (ChIA-PET), like 3C, involves proximity ligation, but also includes a ChIP step (Figure 3B) [103]. ChIA-PET has been applied to diverse targets including the looping interactions mediated by estrogen receptor alpha [103], CTCF-mediated loops [104], RNA polymerase II [105], and a combination of key architectural factors [106].

All of the proximity ligation methods discussed above suffer from the requirement that chromatin be formaldehyde crosslinked, which can distort results [107]. DamID is a method for studying protein–DNA interactions through proximity methylation and although it can be used to map primary structure, such as chromatin-bound proteins [72], it has proved particularly powerful in identifying large-scale domains associated with particular proteins that demarcate nuclear landmarks, such as the nuclear lamina [108] (Figure 3A). A major strength of DamID is that it involves proximity methylation in live cells rather than proximity ligation in fixed chromatin (Figure 3A) [72]. It has been used to identify lamin-associated domains (LADs), which largely comprise late-replicating, transcriptionally inactive chromatin and have significant overlap with inactive TADs [92,108]. TADs have been observed to switch from LAD to non-LAD status during development [93]. The level of nuclear organization probed by DamID and interchromosomal contacts observable by Hi-C or 4C might be viewed as very large-scale tertiary structure or, because they involve multiple chromosomes and chromosome territories, might be likened to quaternary structure by analogy with multimeric proteins. Exciting work in understanding the interactions between the primary and tertiary (or quaternary) structure of the genome has involved dissection of the relationship between nuclear periphery tethering, chromatin state, and transcription, discussed in the next section.

How do the three scales of chromatin architecture interact?

Strong correlative evidence supports interactions between primary and tertiary structure. TAD boundaries are correlated with primary-structure elements such as CTCF-binding sites [13,92], open chromatin, housekeeping and tRNA genes, and high gene density in general [92]. The link between histone modifications and 3D interaction domains is less clear [97]. On the one hand, the borders of histone modification domains often coincide with the borders of TADs and are established by some of the same factors that delineate TADs, such as insulators and highly transcribed genes [91,92,97,109]. TADs can be clustered by epigenetic state [13] and a combined Hi-C/ChIP/RNA-seq

approach has also shown that both gene expression and epigenetic marks that change in response to hormones do so in a consistent manner across TADs [110]. On the other hand, TAD boundaries and chromatin modification boundaries are separable, as illustrated by H3K9me3 domains in differentiated cells: TAD boundaries are present in both pluripotent stem cells and differentiated cells, preceding the appearance of H3K9me3 domains in differentiated cells [92]. Together, these data support the idea that 3D interaction domains can not only coordinate transcription but also partially delimit the diffusion of histone-modifying factors [111].

As a complement to genome-wide correlative data, artificial tethering (reviewed more extensively in [112]) has been primarily studied at the β-globin locus. Ectopic 3D contacts are induced by insertion of insulator elements [113] or by dimerization of the transcription cofactor Ldb1 at an endogenous enhancer in the Locus Control Region of the β-globin locus with an artificial zinc finger–Ldb1 fusion (or a fusion with the Ldb1 self-association domain) targeted to a chosen promoter. These forced contacts induce transcription of the corresponding β-globin gene in the absence of its endogenous activator, GATA1 [114], or even in opposition to the developmental program that normally shuts off transcription of fetal globin genes in adult tissues [115]. Interestingly, the induction of fetal genes in adult cells by forced chromatin contacts was achieved only in human cells and not in mouse, which may be due to a difference in the permissiveness of the chromatin environment [115]. It remains to be seen how broadly results from forced contacts at the β-globin locus will translate to other enhancer–promoter interactions.

A recent paper touches on all three scales, demonstrating that transcriptional activation via a synthetic transcription activator-like effector (TALE)-based TF repositions a gene from the nuclear periphery to the nuclear interior and leads to local chromatin decondensation [116]. By replacing the transcriptional activator domain of the TALE fusion protein with an acidic peptide previously shown to decondense chromatin *in vivo*, the authors functionally separated chromatin decondensation from transcription and showed that although chromatin decondensation is sufficient for nuclear repositioning, it is not sufficient to shift late-replicating chromatin to an earlier replication time, which requires transcription. These results echo the separability of nuclear lamina tethering and transcriptional repression observed in *Caenorhabditis elegans* [117]. The converse experiment, involving forced tethering of reporter loci to the nuclear periphery, has shown that there is likely to be feedback between nuclear positioning and transcription because some (but not all) genes are repressed when relocated to the periphery [118]. Together, these data suggest that the relationships between transcriptional activation, chromatin structure, nuclear organization, and DNA replication are more complex than initially envisioned in classical models of open and closed chromatin.

Concluding remarks: outlook for 3D genomics

Looking forward, several technological developments in the study of primary structure appear ready for exciting applications. First, sequencing technology is being pushed toward longer read lengths that can be used to map the repetitive regions of the genome and structural variation [119]. New bioinformatics methods are also being applied to model repeat structure in centromeres, one of the most fascinating examples of epigenetic modification and specialized chromosome architecture [120].

Second, the refinement of whole-genome TF footprinting assays makes it possible to assay the occupancy of hundreds of TFs in parallel, assuming that the binding motifs of those TFs are known. To that end, the careful quantification of the sequence bias of the nucleases used in footprinting will make this approach more feasible and reliable [121–123]. Footprinting signals due to sequence bias can range over several orders of magnitude, so novel TF binding motifs identified by footprinting should be verified by orthogonal methods and new footprinting methodologies with lower or more manageable sequence bias should be pursued. Sequence bias of nuclease cleavage is also a problem in the currently available maps of nucleosome positions. An alternative method with great potential is chemical mapping of nucleosomes, which uses histone mutants with a unique cysteine on histone H4 to localize a free radical source at the nucleosome dyad and cleave DNA at stereotyped positions around the dyad [124,125]. Chemical mapping has base pair resolution and lacks nuclease-dependent sequence bias but, because of the requirement for a histone mutant, to date has been used only in yeast. Advances in genome editing or novel methods for highly targeted coupling chemistry may make it possible to extend chemical mapping to metazoan genomes, revealing patterns of nucleosome spacing that can provide important clues about higher-order architecture [68,73,124]. Lastly, the quality of primary-structure data from many methods will be vastly improved by precise classification and definition of cell type using single-cell RNA-seq [126] followed by chromatin profiling in well-defined cell populations, which will be likely to require low-input methods (reviewed in [15]). Such low-input primary-structure mapping methods can also reveal the extent to which features like histone modification domains vary from cell to cell or over time. Although these measurements are extremely noisy, useful information can be gleaned from mapping the cell–cell variance in the signal (such as methylation) over the genome and in different cell types [15].

The gap in the multiscale map of genome organization that currently exists at the secondary-structure scale probably must be filled through the development of new methods. Some information may come from 3C family methods as their resolution continues to improve, through innovations like using DNase I digestion to randomly cleave fragments instead of relying on REs [95]. As sequencing costs continue to drop, the fundamental genome coverage and resolution limit of Hi-C will be dictated by the positions of restriction sites, which must be carefully accounted for in Hi-C analyses because they are not random [13,91,127]. DNase I can be used to generate almost arbitrarily short fragments for better genome coverage than in RE-based Hi-C [95]. When combined with polymer modeling, high-resolution proximity ligation methods may be able to discriminate between compaction along the chromatin fiber, as would be the case in regions that adopt a 30-nm fiber, and compaction based on interdigitation of

nucleosomes from a polymer melt [59]. However, proximity ligation alone will probably not provide the base pair or near-base pair resolution that is necessary to infer the 3D architecture.

As the resolution of methods for mapping tertiary chromatin interactions has improved, the challenge of analyzing, visualizing, and interpreting Hi-C data has become more significant. For example, at 5-kb resolution, the contact matrix contains 20 billion pixels and custom parallelized code for graphical processing units was required to analyze these immense data sets [13]. The resulting 2D contact matrices also do not fit the track-based nature of standard genome browsers. To aid in interpreting and visualizing data sets from 3C/Hi-C/ChIA-PET (see below) and related experiments, a software package called Juicebox was created [13], adding to the pre-existing suite of R tools available from a collaboration of groups [128]. In a more generally accessible format, visualization of 3C and related data in a genome track-like format is possible thanks to new features in the WashU Epigenome Browser [129].

Interpretation of 3C and Hi-C interaction frequency maps is complicated by two primary factors. First, proximity that is captured by ligation can arise in multiple ways, including direct protein-mediated association or co-association with the same subnuclear body [130–132]. Second, most proximity ligation experiments average over the chromosome conformations of several million cells and it is unclear whether a Hi-C map represents a series of stable 3D contacts that are present in all cells of a population or the average of stochastic contacts that can vary dramatically between cells. The recently developed single-cell Hi-C [133] showed that single cell contacts were mostly consistent with TADs mapped in ensemble measurements, but interchromosome or inter-TAD contacts varied dramatically between cells. However, single-cell data are by necessity sparse and at very low resolution. A different and potentially more fruitful approach was adopted by another group, which used Monte Carlo modeling of polymer representations of the X chromosome to simulate the possible conformations that are consistent with their 5C data, assuming that proximity ligation frequencies represent contact frequencies in a cell population for any two loci [134]. When compared with 3D FISH (an orthogonal, single-cell method) data, the model successfully predicted the effect of deleting a region at the TAD boundary. This work provides strong evidence that TADs are fluctuating entities present as a compact domain in only a subset of any cell population whose interactions are shaped by transient looping interactions between key loci. Concordantly, TADs mostly disappear in condensed mitotic chromosomes but can be re-established in the next interphase [135]. It was also shown that maintenance of intra-TAD interactions helps to exclude inter-TAD interactions and, using RNA-FISH, that transcriptional activation is not always correlated in a simple way with the transient level of TAD compaction [134].

The emerging picture of the genome's tertiary structure is dynamic, which is unsurprising considering that many *trans*-acting factors undergo fast turnover [136,137]. A 4D view of the genome is therefore needed, with temporal information collected together with high-resolution spatial information both along the chromatin fiber and in 3D interaction space. This acute need is reflected in a recent initiative from the National Institutes of Health (NIH) to fund research on new methods to probe the 4D nucleome [138]. Fluctuations can be studied by surveying variation between cells using single-cell and low-input methods [15,133,139] as well as simulations [134], but temporal information, in both ensemble and single-cell experiments, is necessary to help distinguish causal relationships. Time courses are a powerful and important tool in the study of mechanisms of transcriptional activation, when combined with perturbations of chromatin state and chromatin interactions (Box 1), and have already yielded insights into the nontrivial relationship between enhancer–promoter contacts and transcriptional activation [94,112,114]. At lower throughput but higher temporal resolution, promising approaches include high-resolution imaging of fluorescent proteins tethered to lac or tet operator arrays [140,141], the 'molecular contact memory' method, which uses DamID and an inducible fluorescent fusion protein that specifically binds the methylated DNA [142], and kinetic measurements of chromatin modification and transcription using labeled antibodies and fluorescence microscopy [143].

Box 1. Chromatin perturbation methods

As models are generated to explain the data emerging from chromatin structure mapping experiments performed using the methods described here, a growing number of tools are now available to test these models by actively perturbing chromatin structure and 3D genomic interactions, adding to the classical biologist's toolbox of genetic perturbation in amenable model organisms such as mice, flies, and yeast, the now-established genome editing methods based on CRISPR/Cas9, and RNAi-based knock down of protein factors or RNAs.

- Advances in artificial chromosome design in both humans and yeast suggest that large-scale *in vivo* transcriptional regulation experiments with total control over DNA sequence may be possible in the future [169,170].
- Chromatin modifiers can be ectopically tethered to loci of interest by fusion to DNA-binding domains from the bacterial lac or tet repressor or yeast GAL4, providing their DNA-binding sequences can be inserted into the locus of interest [171–173]. Reversible tethering that uses small-molecule-mediated interaction between the DNA-binding domain and the chromatin modifier has also been implemented to study chromatin modification dynamics [174].
- Transcriptional regulators can be targeted directly to chromatin using domains that recognize particular chromatin modifications, such as H3K27me3 [175].
- The CRISPR/Cas9 system and TAL effector domains have been developed as a platform for designing artificial transcriptional activators or repressors targeted to any sequence [176–178].
- Targeted chromatin decondensation by artificial transcriptional activation or tethering of a chromatin decondensing peptide to induced relocation of a locus from a heterochromatic to a euchromatic compartment of the nucleus shows that perturbations can be achieved at multiple architectural levels simultaneously, due to the multiscale nature of genome regulation [116].
- Artificial zinc-finger proteins have been used to impose forced looping interactions, revealing an early glimpse into the causality relationship between looping and transcriptional activation [114,115].
- The spatial organization of the nucleus can be altered by tethering of chromosome-integrated lac operon arrays to the nuclear periphery by peripheral protein–lacI fusion proteins [118].

The role of noncoding RNAs (ncRNAs) in 3D genome biology, reviewed extensively elsewhere [144,145], is an active area of investigation. For example, recent work has shown that RNAs can help to endow protein *trans*-acting factors with sequence specificity [146]. 3D chromosome architecture can affect ncRNA function [147] and ncRNAs conversely play a role in maintaining 3D interactions [148].

The future of genomics holds exciting opportunities to build on the existing infrastructure of chromatin modification and 3D contact maps by exploring chromatin dynamics and cell–cell or cell-type variation in genome architecture. The interface between genomics and biochemistry also appears ripe for a breakthrough. Although the biochemistry community has expressed skepticism about the usefulness of large-scale epigenetic mapping efforts [149], some vindication may lie in the growing capability for hypothesis-driven methods using classical genetic or biochemical perturbations and tools described in Box 1, as well as in powerful approaches that combine genomics-based identification of factors that regulate genome architecture with *in vitro* characterization of their binding characteristics, such as a series of recent papers that explores the interaction of the PRC2 complex and CTCF with RNAs [150–153]. We hope that such interdisciplinary, multimodal approaches to untangling the puzzle of genomic architecture will serve as archetypes for future work.

Acknowledgments

The authors thank the Straight laboratory and the Greenleaf laboratory for helpful discussions. W.J.G. acknowledges support from NIH grants R21HG007726 and P50HG007735 and the Rita Allen Foundation. V.I.R. acknowledges support from the Walter V. and Idun Berry Postdoctoral Fellowship and the Stanford Center for Systems Biology Seed Grant. The authors apologize to colleagues whose work could not be cited due to space constraints.

Disclaimer statement

W.J.G. is named as an inventor on a patent application filed by Stanford University regarding the ATAC-seq technique and is a scientific Co-founder of Epinomics.

References

1 Lander, E.S. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
2 International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945
3 Luger, K. *et al.* (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* 13, 436–447
4 Lenhard, B. *et al.* (2012) Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat. Rev. Genet.* 13, 233–245
5 Guertin, M.J. and Lis, J.T. (2013) Mechanisms by which transcription factors gain access to target sequence elements in chromatin. *Curr. Opin. Genet. Dev.* 23, 116–123
6 John, S. *et al.* (2011) Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nat. Genet.* 43, 264–268
7 Bell, O. *et al.* (2011) Determinants and dynamics of genome accessibility. *Nat. Rev. Genet.* 12, 554–564
8 Thurman, R.E. *et al.* (2012) The accessible chromatin landscape of the human genome. *Nature* 489, 75–82
9 Shlyueva, D. *et al.* (2014) Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* 15, 272–286
10 Banerji, J. *et al.* (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299–308
11 Buecker, C. and Wysocka, J. (2012) Enhancers as information integration hubs in development: lessons from genomics. *Trends Genet.* 28, 276–284
12 de Wit, E. and de Laat, W. (2012) A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* 26, 11–24
13 Rao, S.S.P. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680
14 Bickmore, W.A. and van Steensel, B. (2013) Genome architecture: domain organization of interphase chromosomes. *Cell* 152, 1270–1284
15 Greenleaf, W.J. (2015) Assaying the epigenome in limited numbers of cells. *Methods* 72, 51–56
16 Krueger, F. *et al.* (2012) DNA methylome analysis using short bisulfite sequencing data. *Nat. Methods* 9, 145–151
17 Gu, H. *et al.* (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* 6, 468–481
18 Smith, Z.D. and Meissner, A. (2013) DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* 14, 204–220
19 Booth, M.J. *et al.* (2013) Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. *Nat. Protoc.* 8, 1841–1851
20 Yu, M. *et al.* (2012) Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat. Protoc.* 7, 2159–2170
21 Zaret, K.S. and Carroll, J.S. (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 25, 2227–2241
22 Giresi, P.G. *et al.* (2007) FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. *Genome Res.* 17, 877–885
23 Hesselberth, J.R. *et al.* (2009) Global mapping of protein–DNA interactions *in vivo* by digital genomic footprinting. *Nat. Methods* 6, 283–289
24 Crawford, G.E. *et al.* (2006) Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). *Genome Res.* 16, 123–131
25 Boyle, A.P. *et al.* (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132, 311–322
26 Buenrostro, J.D. *et al.* (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218
27 Song, L. *et al.* (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res.* 21, 1757–1767
28 Gross, D.S. and Garrard, W.T. (1988) Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57, 159–197
29 Neph, S. *et al.* (2012) An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489, 83–90
30 Neph, S. *et al.* (2012) Circuitry and dynamics of human transcription factor regulatory networks. *Cell* 150, 1274–1286
31 Stergachis, A.B. *et al.* (2014) Conservation of *trans*-acting circuitry during mammalian regulatory evolution. *Nature* 515, 365–370
32 Zentner, G.E. and Henikoff, S. (2014) High-resolution digital profiling of the epigenome. *Nat. Rev. Genet.* 15, 814–827
33 Vierstra, J. *et al.* (2014) Coupling transcription factor occupancy to nucleosome architecture with DNase-FLASH. *Nat. Methods* 11, 66–72
34 Yuan, G-C. *et al.* (2005) Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* 309, 626–630
35 Schones, D.E. *et al.* (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132, 887–898
36 Valouev, A. *et al.* (2011) Determinants of nucleosome organization in primary human cells. *Nature* 474, 516–520
37 Gaffney, D.J. *et al.* (2012) Controls of nucleosome positioning in the human genome. *PLoS Genet.* 8, e1003036
38 Winter, D.R. *et al.* (2013) DNase-seq predicts regions of rotational nucleosome stability across diverse human cell types. *Genome Res.* 23, 1118–1129
39 Park, P.J. (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* 10, 669–680
40 Landt, S.G. *et al.* (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* 22, 1813–1831

41 Thorne, A.W. *et al.* (2004) Native chromatin immunoprecipitation. *Methods Mol. Biol.* 287, 21–44
42 Kasinathan, S. *et al.* (2014) High-resolution mapping of transcription factor binding sites on native chromatin. *Nat. Methods* 11, 203–209
43 Project Consortium, ENCODE. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74
44 Roadmap Epigenomics Consortium *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317–330
45 Kellis, M. *et al.* (2014) Defining functional DNA elements in the human genome. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6131–6138
46 Ernst, J. and Kellis, M. (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* 9, 215–216
47 Hoffman, M.M. *et al.* (2012) Unsupervised pattern discovery in human chromatin structure through genomic segmentation. *Nat. Methods* 9, 473–476
48 Mathelier, A. *et al.* (2014) JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 42, D142–D147
49 Newburger, D.E. and Bulyk, M.L. (2009) UniPROBE: an online database of protein binding microarray data on protein–DNA interactions. *Nucleic Acids Res.* 37, D77–D82
50 Wingender, E. *et al.* (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res.* 24, 238–241
51 Arnold, C.D. *et al.* (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074–1077
52 Zabidi, M.A. *et al.* (2014) Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 518, 556–559
53 Chu, C. *et al.* (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA–chromatin interactions. *Mol. Cell* 44, 667–678
54 Simon, M.D. *et al.* (2013) High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* 504, 465–469
55 Felsenfeld, G. and McGhee, J.D. (1986) Structure of the 30 nm chromatin fiber. *Cell* 44, 375–377
56 Grigoryev, S.A. and Woodcock, C.L. (2012) Chromatin organization – the 30 nm fiber. *Exp. Cell Res.* 318, 1448–1455
57 Li, G. and Reinberg, D. (2011) Chromatin higher-order structures and gene regulation. *Curr. Opin. Genet. Dev.* 21, 175–186
58 Schlick, T. *et al.* (2012) Toward convergence of experimental studies and theoretical modeling of the chromatin fiber. *J. Biol. Chem.* 287, 5183–5191
59 Maeshima, K. *et al.* (2010) Chromatin structure: does the 30-nm fibre exist *in vivo*? *Curr. Opin. Cell Biol.* 22, 291–297
60 Ghirlando, R. and Felsenfeld, G. (2013) Chromatin structure outside and inside the nucleus. *Biopolymers* 99, 225–232
61 Woodcock, C.L. (2006) Chromatin architecture. *Curr. Opin. Struct. Biol.* 16, 213–220
62 Maresca, T.J. and Heald, R. (2006) The long and the short of it: linker histone H1 is required for metaphase chromosome compaction. *Cell Cycle* 5, 589–591
63 Song, F. *et al.* (2014) Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* 344, 376–380
64 Canzio, D. *et al.* (2013) A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature* 496, 377–381
65 Georgel, P.T. *et al.* (2003) Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J. Biol. Chem.* 278, 32181–32188
66 Francis, N.J. *et al.* (2004) Chromatin compaction by a Polycomb group protein complex. *Science* 306, 1574–1577
67 Narlikar, G.J. *et al.* (2013) Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* 154, 490–503
68 Routh, A. *et al.* (2008) Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8872–8877
69 Soboleva, T.A. *et al.* (2012) A unique H2A histone variant occupies the transcriptional start site of active genes. *Nat. Struct. Mol. Biol.* 19, 25–30
70 Robinson, P.J.J. *et al.* (2008) 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J. Mol. Biol.* 381, 816–825
71 Li, G. *et al.* (2010) Highly compacted chromatin formed *in vitro* reflects the dynamics of transcription activation *in vivo*. *Mol. Cell* 38, 41–53
72 van Steensel, B. and Henikoff, S. (2000) Identification of *in vivo* DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat. Biotechnol.* 18, 424–428
73 Koslover, E.F. *et al.* (2010) Local geometry and elasticity in compact chromatin structure. *Biophys. J.* 99, 3941–3950
74 Scheffer, M.P. *et al.* (2011) Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16992–16997
75 Eltsov, M. *et al.* (2008) Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes *in situ*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19732–19737
76 Fussner, E. *et al.* (2012) Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO Rep.* 13, 992–996
77 Eskeland, R. *et al.* (2010) Histone acetylation and the maintenance of chromatin compaction by Polycomb repressive complexes. *Cold Spring Harb. Symp. Quant. Biol.* 75, 71–78
78 Nishino, Y. *et al.* (2012) Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *EMBO J.* 31, 1644–1653
79 Fussner, E. *et al.* (2011) Living without 30 nm chromatin fibers. *Trends Biochem. Sci.* 36, 1–6
80 Collepardo-Guevara, R. and Schlick, T. (2014) Chromatin fiber polymorphism triggered by variations of DNA linker lengths. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8061–8066
81 Wu, C. *et al.* (2007) A variable topology for the 30-nm chromatin fibre. *EMBO Rep.* 8, 1129–1134
82 Horowitz, R.A. *et al.* (1994) The three-dimensional architecture of chromatin *in situ*: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. *J. Cell Biol.* 125, 1–10
83 Dekker, J. *et al.* (2002) Capturing chromosome conformation. *Science* 295, 1306–1311
84 de Laat, W. and Dekker, J. (2012) 3C-based technologies to study the shape of the genome. *Methods* 58, 189–191
85 Tolhuis, B. *et al.* (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10, 1453–1465
86 Palstra, R-J. *et al.* (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* 35, 190–194
87 Simonis, M. *et al.* (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38, 1348–1354
88 Zhao, Z. *et al.* (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* 38, 1341–1347
89 Dostie, J. *et al.* (2006) Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 16, 1299–1309
90 Hughes, J.R. *et al.* (2014) Analysis of hundreds of *cis*-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* 46, 205–212
91 Lieberman-Aiden, E. *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293
92 Dixon, J.R. *et al.* (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380
93 Nora, E.P. *et al.* (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385
94 Jin, F. *et al.* (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290–294
95 Ma, W. *et al.* (2015) Fine-scale chromatin interaction maps reveal the *cis*-regulatory landscape of human lincRNA genes. *Nat. Methods* 12, 71–78
96 Yang, J. and Corces, V.G. (2012) Insulators, long-range interactions, and genome function. *Curr. Opin. Genet. Dev.* 22, 86–92
97 Van Bortle, K. and Corces, V.G. (2013) The role of chromatin insulators in nuclear architecture and genome function. *Curr. Opin. Genet. Dev.* 23, 212–218
98 Phillips, J.E. and Corces, V.G. (2009) CTCF: master weaver of the genome. *Cell* 137, 1194–1211

99 Pope, B.D. *et al.* (2014) Topologically associating domains are stable units of replication-timing regulation. *Nature* 515, 402–405
100 Phillips-Cremins, J.E. *et al.* (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153, 1281–1295
101 Kornberg, R.D. (2005) Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* 30, 235–239
102 Montavon, T. *et al.* (2011) A regulatory archipelago controls Hox genes transcription in digits. *Cell* 147, 1132–1145
103 Fullwood, M.J. *et al.* (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* 462, 58–64
104 Handoko, L. *et al.* (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* 43, 630–638
105 Li, G. *et al.* (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148, 84–98
106 Heidari, N. *et al.* (2014) Genome-wide map of regulatory interactions in the human genome. *Genome Res.* 24, 1905–1917
107 Gavrilov, A.A. *et al.* (2013) Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. *Nucleic Acids Res.* 41, 3563–3575
108 Guelen, L. *et al.* (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948–951
109 Sexton, T. *et al.* (2012) Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148, 458–472
110 Le Dily, F. *et al.* (2014) Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* 28, 2151–2162
111 Nora, E.P. *et al.* (2013) Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? *Bioessays* 35, 818–828
112 Deng, W. and Blobel, G.A. (2014) Manipulating nuclear architecture. *Curr. Opin. Genet. Dev.* 25, 1–7
113 Hou, C. *et al.* (2008) CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 20398–20403
114 Deng, W. *et al.* (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149, 1233–1244
115 Deng, W. *et al.* (2014) Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* 158, 849–860
116 Therizols, P. *et al.* (2014) Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science* 346, 1238–1242
117 Towbin, B.D. *et al.* (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* 150, 934–947
118 Finlan, L.E. *et al.* (2008) Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet.* 4, e1000039
119 Chaisson, M.J.P. *et al.* (2014) Resolving the complexity of the human genome using single-molecule sequencing. *Nature* 517, 608–611
120 Miga, K.H. *et al.* (2014) Centromere reference models for human chromosomes X and Y satellite arrays. *Genome Res.* 24, 697–707
121 He, H.H. *et al.* (2012) Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. *Genome Res.* 22, 1015–1025
122 Meyer, C.A. and Liu, X.S. (2014) Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nat. Rev. Genet.* 15, 709–721
123 Sung, M-H. *et al.* (2014) DNase footprint signatures are dictated by factor dynamics and DNA sequence. *Mol. Cell* 56, 275–285
124 Brogaard, K. *et al.* (2012) A map of nucleosome positions in yeast at base-pair resolution. *Nature* 486, 496–501
125 Moyle-Heyrman, G. *et al.* (2013) Chemical map of *Schizosaccharomyces pombe* reveals species-specific features in nucleosome positioning. *Proc. Natl. Acad. Sci. U.S.A.* 110, 20158–20163
126 Patel, A.P. *et al.* (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344, 1396–1401
127 Imakaev, M. *et al.* (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods* 9, 999–1003
128 Servant, N. *et al.* (2012) HiTC: exploration of high-throughput 'C' experiments. *Bioinformatics* 28, 2843–2844
129 Zhou, X. *et al.* (2013) Exploring long-range genome interactions using the WashU Epigenome Browser. *Nat. Methods* 10, 375–376
130 Dekker, J. *et al.* (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev. Genet.* 14, 390–403
131 Belmont, A.S. (2014) Large-scale chromatin organization: the good, the surprising, and the still perplexing. *Curr. Opin. Cell Biol.* 26, 69–78
132 Williamson, I. *et al.* (2014) Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence *in situ* hybridization. *Genes Dev.* 28, 2778–2791
133 Nagano, T. *et al.* (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59–64
134 Giorgetti, L. *et al.* (2014) Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* 157, 950–963
135 Naumova, N. *et al.* (2013) Organization of the mitotic chromosome. *Science* 342, 948–953
136 van Werven, F.J. *et al.* (2009) Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. *Nat. Struct. Mol. Biol.* 16, 1043–1048
137 Cheutin, T. *et al.* (2003) Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* 299, 721–725
138 Office of the Director, N.I.H. (2014) *NIH Programs to Focus on Emerging Areas of Science*, NIH
139 Trombetta, J.J. *et al.* (2014) Preparation of single-cell RNA-seq libraries for next generation sequencing. *Curr. Protoc. Mol. Biol.* 107, 4.22.1–4.22.17
140 Masui, O. *et al.* (2011) Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. *Cell* 145, 447–458
141 Belmont, A.S. *et al.* (2010) Insights into interphase large-scale chromatin structure from analysis of engineered chromosome regions. *Cold Spring Harb. Symp. Quant. Biol.* 75, 453–460
142 Kind, J. *et al.* (2013) Single-cell dynamics of genome–nuclear lamina interactions. *Cell* 153, 178–192
143 Stasevich, T.J. *et al.* (2014) Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature* 516, 272–275
144 Rinn, J. and Guttman, M. (2014) RNA Function. RNA and dynamic nuclear organization. *Science* 345, 1240–1241
145 Cech, T.R. and Steitz, J.A. (2014) The noncoding RNA revolution – trashing old rules to forge new ones. *Cell* 157, 77–94
146 Batista, P.J. and Chang, H.Y. (2013) Long noncoding RNAs: cellular address codes in development and disease. *Cell* 152, 1298–1307
147 Engreitz, J.M. *et al.* (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341, 1237973
148 Hacisuleyman, E. *et al.* (2014) Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* 21, 198–206
149 Madhani, H.D. *et al.* (2008) Epigenomics: a roadmap, but to where? *Science* 322, 43–44
150 Davidovich, C. *et al.* (2013) Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. Mol. Biol.* 20, 1250–1257
151 Davidovich, C. *et al.* (2015) Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA. *Mol. Cell* 57, 552–558
152 Kung, J.T. *et al.* (2015) Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell* 57, 361–375
153 Cifuentes-Rojas, C. *et al.* (2014) Regulatory interactions between RNA and Polycomb repressive complex 2. *Mol. Cell* 55, 171–185
154 Auerbach, R.K. *et al.* (2009) Mapping accessible chromatin regions using Sono-Seq. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14926–14931
155 Stergachis, A.B. *et al.* (2013) Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. *Cell* 154, 888–903
156 Song, L. and Crawford, G.E. (2010) DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb. Protoc.* 2010, http://dx.doi.org/10.1101/pdb.prot5384
157 Henikoff, J.G. *et al.* (2011) Epigenome characterization at single base-pair resolution. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18318–18323

158 Cuddapah, S. *et al.* (2009) Native chromatin preparation and Illumina/Solexa library construction. *Cold Spring Harb. Protoc.* 2009, http://dx.doi.org/10.1101/pdb.prot5237
159 Brind'Amour, J. *et al.* (2015) An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* 6, 6033
160 Rhee, H.S. and Pugh, B.F. (2012) ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy. *Curr. Protoc. Mol. Biol.* 100, 21.24.1–21.24.14
161 Hagège, H. *et al.* (2007) Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat. Protoc.* 2, 1722–1733
162 Splinter, E. *et al.* (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* 20, 2349–2354
163 Noordermeer, D. *et al.* (2011) The dynamic architecture of Hox gene clusters. *Science* 334, 222–225
164 Stadhouders, R. *et al.* (2013) Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions. *Nat. Protoc.* 8, 509–524
165 Sanyal, A. *et al.* (2012) The long-range interaction landscape of gene promoters. *Nature* 489, 109–113
166 van Steensel, B. and Dekker, J. (2010) Genomics tools for unraveling chromosome architecture. *Nat. Biotechnol.* 28, 1089–1095
167 Vogel, M.J. *et al.* (2007) Detection of *in vivo* protein–DNA interactions using DamID in mammalian cells. *Nat. Protoc.* 2, 1467–1478
168 Tolhuis, B. *et al.* (2006) Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila* melanogaster. *Nat. Genet.* 38, 694–699
169 Bergmann, J.H. *et al.* (2012) HACking the centromere chromatin code: insights from human artificial chromosomes. *Chromosome Res.* 20, 505–519
170 Annaluru, N. *et al.* (2014) Total synthesis of a functional designer eukaryotic chromosome. *Science* 344, 55–58
171 Li, Y. *et al.* (2003) Effects of tethering HP1 to euchromatic regions of the *Drosophila* genome. *Development* 130, 1817–1824
172 Kagansky, A. *et al.* (2009) Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. *Science* 324, 1716–1719
173 Burgess, R.C. *et al.* (2014) Activation of DNA damage response signaling by condensed chromatin. *Cell Rep.* 9, 1703–1717
174 Hathaway, N.A. *et al.* (2012) Dynamics and memory of heterochromatin in living cells. *Cell* 149, 1447–1460
175 Haynes, K.A. and Silver, P.A. (2011) Synthetic reversal of epigenetic silencing. *J. Biol. Chem.* 286, 27176–27182
176 Maeder, M.L. *et al.* (2013) Robust, synergistic regulation of human gene expression using TALE activators. *Nat. Methods* 10, 243–245
177 Gilbert, L.A. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451
178 Konermann, S. *et al.* (2015) Genome-scale transcriptional activation by an engineered CRISPR–Cas9 complex. *Nature* 517, 583–588