The impact of space and time on the functional output of the nucleus

Marcelo Nollmann, Isma Bennabi, Markus Götz, Thomas Gregor

To cite this version:


HAL Id: pasteur-03216312
https://hal-pasteur.archives-ouvertes.fr/pasteur-03216312
Submitted on 13 Oct 2021

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

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

Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International License
The impact of space and time on the functional output of the nucleus

Marcelo Nollmann¹, Isma Bennabi², Markus Götz¹, Thomas Gregor²,³,*

¹Centre de Biologie Structurale, CNRS UMR5048, INSERM U1054, Univ Montpellier, 60 rue de Navacelles, 34090, Montpellier, France

²Department of Stem Cell and Developmental Biology, CNRS UMR3738, Institut Pasteur, 25 rue du docteur Roux, 75015 Paris, France

³Joseph Henry Laboratory of Physics & Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

*correspondence: thomas.gregor@pasteur.fr

Abstract

Over the past two decades, it has become clear that the multi-scale spatial and temporal organization of the genome has important implications for nuclear function. This chapter centers on insights gained from recent advances in light microscopy on our understanding of transcription. Particular focus is given to the genomic scales where most functional interactions occur, namely kb–Mb. We discuss the relevant spatial and temporal scales that shape nuclear order and their consequences on regulatory components and function. The emerging picture is that spatiotemporal constraints increase the complexity in transcriptional regulation. This brings new challenges to the fore, such as uncertainty about how information travels from factors through the genome and space to generate a functional output.

Short title: Dynamics and function of the nucleus
Introduction

There is a growing appreciation that gene function is connected to the dynamic structure of chromosomes. In particular, spatiotemporal aspects of genome architecture are crucial to our understanding of eukaryotic gene expression, and thus to the main functional output of the nucleus (van Steensel and Furlong 2019). However, the mechanistic underpinnings and actual causal links between structure and function are scarce and present an obvious challenge for the next decade. Generation of vast atlases of Hi-C proximity maps over the past ten years provided a starting point for such studies (Dekker 2016). They underscore the existence of DNA organization at the kb-Mb scales as an inherent, functionally important, component in gene activation. However, these atlases are limited in their ability to link single-cell three-dimensional genome structures to specific transcriptional states. In addition, a characterization of how dynamic long-range interactions regulate single-cell transcriptional dynamics is missing, especially at the genomic scales where cis-regulatory elements mediate functional interactions.

Cis-regulatory elements (CREs) are non-coding DNA regions that regulate transcription. They include enhancers, chromatin insulators, silencers, and promoters (Wittkopp and Kalay 2011). Enhancers are short regulatory DNA sequences that control gene activity and provide the genetic implementation for the dynamic control of gene expression (de Laat and Duboule 2013; Bolt and Duboule 2020). The precise binding of regulators to promoter or enhancer sequences (Fig. 1A,B) gives rise to the time and tissue-specific activation of subsets of genes to confer cell identity. Since their discovery four decades ago (Banerji et al. 1981; Mercola et al. 1983; Gillies et al. 1983; Banerji et al. 1983; Moreau et al. 1981), enhancers have been largely regarded as autonomous, modular units, capable of activating transcription in a location and orientation independent manner. Enhancers are often at large distances from their respective target-gene promoters, most times with additional non-target genes found within the intervening sequences (Furlong and Levine 2018; Schoenfelder and Fraser 2019). The canonical model is that enhancers regulate transcription by physically interacting with promoters over large genomic distances (Fig. 1C) (Blackwood and Kadonaga 1998). Whole-genome methods have shown that the human genome is riddled with enhancers, with estimates ranging from hundreds of thousands to over a million (Schoenfelder and Fraser 2019; Xu et al. 2020; Pennacchio et al. 2013). On average, a typical human gene is regulated by at least 10–20 different enhancers (Sanyal et al. 2012), raising the possibility that multiple enhancers may physically contact the same promoter (Fig. 1D,E).

Remarkably, this network of interactions may be modulated by other CREs, such as chromatin insulators. These are short, cis-regulatory sequences that block communication between promoters and enhancers (Reitman et al. 1990; Geyer and Corces 1992; Cai and Levine 1995). Thousands of sites genome-wide are characterized as insulators, particularly enriched in intergenic and promoter regions. Insulators play a role in the formation of long-range interactions (Vogelmann et al. 2011; Yang and Corces 2012; Vogelmann et al. 2014) and are involved in the global regulation of transcription (Bushey et al. 2009). Chromosome conformation capture (3C)-based assays (Lieberman-Aiden et al. 2009) revealed the existence of well-defined kb–Mb genomic regions displaying locally enhanced chromatin interactions (Sexton et al. 2012; Dixon et al. 2012; Nora et al. 2012). These genomic regions, typically called topologically associating domains (TADs), tend to be demarcated by chromatin insulators (Rao et al. 2015; Sexton et al. 2012; Hou et al. 2012) and often encapsulate enhancers and their target genes (Shen et al. 2012; Symmons et al. 2014; Neems et al. 2016; Ji et al. 2016; Dowen et al. 2014; Ron et al. 2017) (Fig. 1E).
Presently, the role of TADs in gene activity is still an open question (for a review see (Cavalheiro et al. 2021; Schoenfelder and Fraser 2019)).

Disruption of TAD architecture, either by local duplications, deletions, or inversions that fuse adjacent TADs or form new TADs, were reported to trigger developmental defects in mammals because of improper enhancer–promoter interactions (Lupiáñez et al. 2015; Northcott et al. 2014; Franke et al. 2016). In contrast, global disruption of TADs by CTCF or cohesin depletion caused only mild changes in gene expression (Nora et al. 2017; Rao et al. 2017; Schwarzer et al. 2017), suggesting that TADs may be important for regulating only a subset of genes. The reason for these apparent discrepancies may arise from the intrinsic dynamic properties of the chromatin fiber and transcription processes. Most studies assessing the transcriptional roles of TADs stem from a qualitative view: ensemble, static measurements of TAD borders or contacts between CREs and promoters, combined with the assessment of bulk transcriptional outputs. This view entirely lacks the temporal and spatial resolutions that are critical to understanding the genesis of these intrinsically dynamic processes, or to address key questions about the roles of TADs in transcriptional regulation: Are chromosome dynamics random and thus a key component to the stochasticity of gene expression? How can remote enhancers direct the correct spatial and temporal control of transcription? How do multiple enhancers interact dynamically to control promoter accessibility?

Renewed focus on spatial nuclear chromatin organization and its potential impact on gene regulation has brought a new twist to our understanding of eukaryotic transcription, broadening the potential angles evolution can use to interfere with the underlying mechanisms. Over the past forty years, we have gone from transcription factors and
machinery binding directly to promoters, to single or multiple enhancers, to mediated transcriptional control, to distal enhancer–promoter communication over Mb in cis and even in trans (Figure 1). Now a novel, spatial component comes into play, where entire loci are organized in three dimensions, and where organization can have functional control. Recently, imaging technologies provided insightful tools to probe this organization and its dynamic impact on transcription. Here, we review how these new tools are starting to provide new insights into otherwise inaccessible fundamental biological questions. In particular, understanding how spatiotemporal changes in chromosome structure modulate genome function requires progress at four levels: 1) the relationship between genomic and physical scales, 2) the physical organization of regulatory elements, 3) the time scales involved in chromosome organization and transcription, and 4) the interplay between the spatiotemporal dynamics of the genome and its function. To each of these levels, we devoted one of the following sections.

1. Spatial scales in the nucleus
At the heart of the control of eukaryotic gene regulation is the organization and function of transcriptional enhancers, the major constituents of the noncoding genome controlling gene activity (Furlong and Levine 2018). A typical gene is regulated by multiple different enhancers. Many of these enhancers possess overlapping regulatory activities, raising questions about proper cis-regulatory “trafficking” whereby the correct enhancers interact with the appropriate target promoters (Chopra et al. 2009; Bushey et al. 2008). What is the influence of the exact position and affinity of DNA binding factors to cis-regulatory elements on transcriptional output? Answering these questions requires not only the relative genomic positioning of CREs (1D) (Negre et al. 2011; ENCODE Project Consortium 2012), but also the direct detection of enhancer–promoter interactions in 3D to understand the impact of spatial chromatin architecture on function.

1D to 3D: unintuitive expectations
The physical properties of the chromatin fiber can be modeled using conventional polymer theory. In the simplest case, chromatin can be approximated by a polymer composed of a chain of identical monomers (Nelson 2003; Rippe 2001; de Gennes and Gennes 1979; I?U. Grosberg et al. 2011; Doi 1996; Wiggins et al. 2006). The dynamics of the polymer chain are then governed by thermal fluctuations, excluded volume interactions, and rigidity parameters such as the effective persistence length. This simple model has been successfully used to predict the expected mean three dimensional (3D) distance between two DNA loci as a function of their genomic distance in sequence space (1D) (Rosa and Everaers 2008). Importantly, the conversion between 1D and 3D distances follows a power-law (Mirny 2011) with a fractional exponent that varies between species, mainly due to changes in the physical properties of chromatin, such as the persistence length, the molecular composition of the chromatin fiber, genomic sizes, and nuclear volumes (Mirny 2011).

Importantly, the identified power-law relation between distances in sequence (1D) and in physical (3D) space is non-linear. For instance, so-called distal enhancers in Drosophila are typically found at relatively short genomic distances (1-100 kb) (Ghavi-Helm et al. 2014) that can translate into relatively large average physical distances (200-400 nm) (Rosa and Everaers 2008; Cardozo Gizzi et al. 2019). Similarly, in mammals, enhancers found hundreds of kb from their target promoter are on average very far away (> 1μm (Rosa and Everaers 2008)). However, chromosomes are confined within the limited nuclear space, thus DNA loci located very far in genomic space (e.g. ~2 Mbp), or even in other chromosomes, are not proportionally that far in 3D space (e.g. 1 μm (Cattoni et al. 2017)).
1D sequence encodes protein binding to specific sites, which in turn can modulate the folding of chromatin in 3D. For instance, the binding of chromatin insulators (e.g. CTCF), transcription factors, heterochromatin-associated complexes (e.g. Polycomb, HP1), or architectural proteins (e.g. cohesin, condensin) (Rowley and Corces 2016) promote 3D bridges between chromatin regions that can considerably affect the frequencies of interactions expected for a homopolymer (Rowley and Corces 2016). For example, 3D loops between converging CTCF sites can produce specific 3D interactions between TAD borders (Rao et al. 2015) while cohesin-mediated loops can affect the contact frequencies within TADs (Rao et al. 2017; Nora et al. 2017). Alternatively, chromatin hubs linking multiple genomic loci can lead to the spatial clustering of regulatory elements (Allahyar et al. 2018; Oudelaar et al. 2018; Beagrie et al. 2017; Quinodoz et al. 2018; Espinola et al. 2021). Thus, both passive binding (e.g. CTCF, transcription factors) and active processes (e.g. cohesin/condensin looping, transcriptional elongation) can alter 3D chromatin organization at the kilobase (kb) to megabase (Mb) scales, bringing into close spatial proximity loci that would be expected to reside far apart in a pure homopolymeric chromosome. Understanding these processes and their functional consequences holds potentially far-reaching insights into genome evolution, and requires methods able to dissect chromatin interactions in 3D.

3D mapping: sequencing and imaging approaches
Chromatin conformation can be measured by intrinsically different classes of methods. One class is sequencing based, such as 3C and its derivatives, where chromatin contact frequencies measure the proximity between genomic loci averaged over a population of cells. The range over which genomic loci are crosslinked by ligation is debated; computational models estimate the distance to be around 100nm (Giorgetti et al. 2014).

An alternative class is based on microscopy techniques, such as fluorescence in situ hybridization (DNA-FISH), that measure pairwise distances directly. These are used to estimate imaging-derived contact frequencies by computing the proportion of cells displaying pairwise distances smaller than a critical radius $R_{M}$ (equivalent to the spherical pairwise distance distribution up to $R_{M}$) (Cattoni et al. 2017; Finn et al. 2019; Mateo et al. 2019; Bintu et al. 2018; Wang et al. 2016; Cardozo Gizzi et al. 2019).

It is worth noting that in neither case the notion of ‘contact’ implies physical interaction, but rather an estimate of whether two genomic loci are spatially close to each other. Thus, in the remainder of this chapter, we will refer to ‘proximity’ rather than ‘contact frequency’. $R_{ur}$-values can be derived from control experiments where a single genomic locus is imaged in multiple colors/cycles (Cattoni et al. 2017; Cardozo Gizzi et al. 2019; Mateo et al. 2019). $R_{ur}$-values between 150–500 nm were typically used, and produced good correlations between Hi-C and microscopy-based proximity frequencies (Wang et al. 2016; Bintu et al. 2018; Cardozo Gizzi et al. 2019; Mateo et al. 2019; Su et al. 2020). As such, the working definition of proximity is different for genomic-based and imaging-based methods, and its exact meaning will likely shift in future measurements with increasing genomic and optical resolutions. Determining the true measure for each method will likely involve a correlation analysis that reveals the scales at which proximity frequencies are correlated for both methods.

In the past, it has become common practice to assess the specificity of 3D interactions by using DNA-FISH to compare the mean 3D pairwise distance between two candidate genomic loci with that of a control that resides at the same genomic distance (Rao et al. 2015; Benabdallah et al. 2019; Ogiyama et al. 2018). However, relying purely on mean distances
can be deceptive in cases where the full pairwise distance distribution departs from that of a single species or when specific 3D interactions are rare. For instance, while proximity frequency and mean spatial distance are inversely correlated for a homopolymer, this is not necessarily the case in presence of sequence-dependent 3D interactions, where a 3D loop is rather expected to lead to a bi-modal distance distribution (Fudenberg and Imakaev 2017; Giorgetti et al. 2014). In this case, the mean pairwise distance fails to capture sequence-specific looping interactions, particularly if these occur at low frequencies (e.g. < 10%).

Even more counterintuitively, an increase in the frequency of sequence-specific interactions can lead to lower, equal, or even higher mean pairwise distances depending on the shape of the pairwise distance distribution (Giorgetti and Heard 2016; Fudenberg and Imakaev 2017). Thus, in a general case, a change in the mean pairwise distance between two genomic regions may not necessarily reflect changes in proximity. Instead, the proximity measured from the full pairwise distance distribution obtained from a DNA-FISH experiment should be preferred to assess changes in 3D chromatin organization (see above). These considerations are critical to derive models of enhancer function from DNA-FISH experiments.

Heterogeneity in chromosome organization is well-documented by the width of pairwise distance distributions. The variances of these distributions are often comparable to, and scale linearly with the mean pairwise distance (Giorgetti et al. 2014; Cattoni et al. 2017). These large single-cell heterogeneities in pairwise distances (Finn et al. 2019; Cattoni et al. 2017; Giorgetti et al. 2014; Bintu et al. 2018) are consistent with large cell-to-cell variations in TAD volumes (Nir et al. 2018; Szabo et al. 2018; Luppino et al. 2020; Boettiger et al. 2016). These structural heterogeneities can originate from several sources. First, they can arise from the highly dynamic behaviour of the chromatin fiber at multiple spatial and temporal scales (for a review see (Tortora et al. 2020)). Second, heavy modulation of nuclear morphology during the cell cycle. Third, active DNA management processes acting on chromatin, such as transcription and replication, are intrinsically dynamic and their status varies between single cells. Finally, DNA binding proteins such as insulators (e.g. CTCF) (Hansen et al. 2017) or transcription factors (Normanno et al. 2015; Izeddin et al. 2014) can rapidly bind to and dissociate from DNA (seconds to minutes time scales) to shape the dynamics of local and global chromatin conformation. As a consequence, it is perhaps not surprising that enhancer–promoter interactions are highly heterogeneous and dynamic in single cells (see sections 3-4 below).
molecular models (1–50 nm).

Thus, there is a notable gap between genomic scales from in vivo enhancer–promoter distance measurements (200–350 nm) and molecular scales from atomic-resolution molecular models (1–50 nm).

Bridging the genomic and molecular scales

Genomic and imaging methods can map chromatin proximities in the 100–500 nm range, however the factors mediating these interactions can be much smaller. DNA binding domains have typical sizes in the sub-nm range (e.g. CTCF zinc fingers 4-8, Fig. 2A) (Yin et al. 2017), similarly to protein-protein interaction domains (e.g. SA2-SCC1 cohesin subunit bound to CTCF N-terminal domain, Fig. 2B) (Li et al. 2020b). DNA-bound transcription factor sizes are commonly in the nanometer range (e.g. lactose operon repressor bound to operator DNA, Fig. 2C) (Lewis et al. 1996), while distances between DNA segments bridged by single protein complexes can be ~10 nm apart (e.g. gamma-delta resolvase bound to site I DNA, Fig. 2D) (Li et al. 2005). Similarly, nucleosomes are ~10 nm in size (Fig. 2E) (Tachiwana et al. 2010). Multi-subunit machines, such as the Mediator complex, RNA polymerase II, or the SWI/SNF chromatin remodeling complex can reach sizes between 14–22 nm (Fig. 2F-H) (Robinson et al. 2015; Nozawa et al. 2017; Armache et al. 2005; Wagner et al. 2020), while condensins can be up to 40 nm long (e.g. condensin complex from S. cerevisiae, Fig. 2I) (Lee et al. 2020).

It is generally accepted that enhancer function requires looping to the promoter (reviewed in Bulger and Groudine 2011; Schwarzer and Spitz 2014). Indeed, imaging-based studies recently reported that transcriptional activation can occur when enhancers get closer than 200–350 nm to promoters (Chen et al. 2018; Li et al. 2020a). Similarly, 3C data combined with modeling described the formation of 100–200 nm active ‘cages’ (Di Stefano et al. 2020). Thus, there is a notable gap between genomic scales from in vivo enhancer–promoter distance measurements (200–350 nm) and molecular scales from atomic-resolution molecular models (1–50 nm).
Part of the explanation for this discrepancy may be explained by the formation of higher-order structures. In fact, multiple factors involved in transcriptional regulation and chromosome organization possess low-complexity/intrinsically disordered regions (IDRs) that may attain considerably large physical distances (Watson and Stott 2019). For instance, histone tails constitute sites of extensive post-translational modifications that encode epigenetic information, and transcription factors have been long predicted to encode extensive IDRs (Liu et al. 2006). IDRs can also fold by interacting with other factors, such as a region of the intrinsically disordered N-terminal domain of CTCF when in interaction with the SA2-SCC1 subunit of cohesin (Li et al. 2020b). Notably, these interactions are necessary for loop formation (Li et al. 2020b; Pugacheva et al. 2020). Finally, IDRs can play many functions (Watson and Stott 2019), including formation of phase-separated compartments (Chong et al. 2018; Sabari et al. 2018; Boija et al. 2018). These alternative models will be explored in the following section.

2. Spatial clustering of regulatory components

Gene expression is a highly regulated process that relies on a multitude of interactions between cis- and trans-acting factors. While CREs are cis-regulatory genetic elements (such as enhancers, promoters, insulators, and silencers), trans-factors are molecular components, such as proteins, protein complexes and non-coding RNAs that are involved in gene regulation in the form of TFs, coactivators (e.g. the Mediator complex), RNA Pol II, IncRNA, etc (Schoenfelder and Fraser 2019). At present, very limited knowledge exists about the relative spatial organization and interactions of CREs and trans-factors in the nucleus. Which of these components, and how many, come together in the nuclear space? Do they display distinct spatial organizations? How are interactions between these elements mediated?

Recent findings revealed that cis- and trans- regulatory components can come into close spatial proximity, and hinted at a number of physical mechanisms that may be implicated in this process. No common terminology for the observed structures has been found up until now (A and Weber 2019). Thus, we made the choice to refer to clusters of chromatin regions containing CREs as CRE-hubs, and nuclear aggregates of proteins or protein complexes as foci, without implying an underlying physical mechanism for either of them (Fig. 3A). The following sections will summarize evidence for the formation of CRE-hubs, and transcription-associated foci, and the physical mechanisms that could describe their formation.

![Figure 3. CRE-hubs, protein foci, and phase separated condensates. (A) Nuclear protein foci may or may not be DNA associated/bound. The relationship between protein foci and CRE-hubs is so far unknown. (B) Model for transcription-associated foci mediating enhancer–promoter interactions. (C) Two mechanisms of condensate](image-url)
formation via phase separation. Liquid-liquid phase separation (LLPS) is driven by weak, multivalent interactions between the constituents of the condensate (red dots) and does not require a scaffolding polymer. Conversely, polymer-polymer phase separation (PPPS) results from crosslinking the chromatin scaffold by a bridging factor (red rectangles).

**Hubs of cis-regulatory elements**

In the canonical model, physical interaction between an enhancer and a promoter is required for enhancer action. The typical eukaryotic gene is regulated by multiple enhancers, particularly during development (Osterwalder et al. 2018; Fulco et al. 2019; Oudelaar and Higgs 2020), and many enhancers may be shared by multiple target genes (Ghavi-Helm et al. 2014), sometimes even simultaneously (Fukaya et al. 2016). These observations suggest that multiple enhancers and promoters come into close spatial proximity during transcription.

To directly probe the spatial clustering of multiple CREs, standard 3C-based techniques are not sufficient as they detect binary chromatin interactions. Consequently, sequencing and imaging-based technologies have been developed to detect multi-way interactions. Amongst the former, 'genome architecture mapping' (GAM) (Beagrie et al. 2017) and 'split-pool recognition of interactions by tag extension' (SPRITE) (Quinodoz et al. 2018) were able to detect three-way interactions between super-enhancers in fixed mouse embryonic stem cells (mESCs). Alternatively, proximity-ligation methods were used to describe the existence of multi-way enhancer interactions at the β- and α-globin locus during cell differentiation (Oudelaar et al. 2018; Allahyar et al. 2018). All in all, these results showed that multiple CREs can be captured in close proximity (Oudelaar et al. 2018; Allahyar et al. 2018). However, these methods are unable to detect multiple chromatin contacts and transcriptional status at once, thus it was unclear whether clustering of CREs has a functional role.

Multiplexed imaging methods are ideally suited to tackle this question. Several complementary approaches rely on similar principles: ‘chromatin tracing’ (Bintu et al. 2018; Wang et al. 2016), 'Hi-M' (Cardozo Gizzi et al. 2019), 'optical reconstruction of chromatin architecture' (ORCA) (Mateo et al. 2019), ‘chromosome walking’ (Nir et al. 2018), as well as ‘seqFISH+’ (Eng et al. 2019) and ‘MINA’ (Liu et al. 2020). They combine microfluidics, wide-field microscopy, and Oligopaint FISH (Beliveau et al. 2012) to resolve the physical 3D position of tens to thousands of genomic loci in the nucleus, and can reach kilobase and nanometric precision.

These multiplexed imaging methods were recently used to simultaneously determine chromosome organization and transcription. ORCA revealed that the architecture of a Hox Polycomb TAD changes between different cell-types in *Drosophila* embryos ((Mateo et al. 2019)). Hi-M was applied to detect CRE interactions and transcriptional status during early *Drosophila* development (Espinola et al. 2021). This study revealed that CRE spatial clustering pre-dates gene activation and does not seem to depend on transcriptional state, with transcriptionally active and silent cells displaying similar CRE hubs (Espinola et al. 2021). These results are consistent with a concurrent Hi-C study demonstrating that chromatin conformation is independent of gene regulation (Ing-Simmons et al. 2020), and with earlier reports showing that enhancer–promoter proximity can precede gene activation in mice or during *Drosophila* embryonic development (Paliou et al. 2019; Montavon et al. 2011; Ghavi-Helm et al. 2014). However, it is important to bear in mind that cell-specific changes in enhancer–promoter interaction networks during differentiation have also been well documented (reviewed in (Schoenfelder and Fraser 2019)). Thus, we hypothesize that the roles of CRE-hubs may be regulated either by altering their 3D structure and/or by modifying the cocktail of trans-acting factors they are bound by. These changes could be realised in different cell types by fine-tuning the abundance and binding of trans-acting factors.
Transcription-associated nuclear foci

In eukaryotes, CREs are bound by TFs and coactivators (e.g. the histone acetyltransferase p300 or the Mediator complex). Interestingly, conventional and super-resolved imaging approaches revealed that many of these factors form nuclear foci (Liu et al. 2014; Dufourt et al. 2018; Boija et al. 2018; Chong et al. 2018; Mir et al. 2018; Cho et al. 2018; Sabari et al. 2018; Tsai et al. 2017; Cisse et al. 2013). This observation raised the possibility that trans-acting factors may form nuclear micro-environments where genes are co-regulated. In some instances, transcription-associated nuclear foci have been reported to contain mRNAs (Boija et al. 2018; Sabari et al. 2018) and to associate to chromatin (Chong et al. 2018; Sabari et al. 2018; Cho et al. 2018), consistent with the observation of CRE-hubs (Figs. 3A-B). However, it is not clear whether this is the norm or the exception.

The functional roles of transcription-associated nuclear foci in gene regulation are still unclear. For instance, are CREs always associated with nuclear foci? How many CREs take part in the formation of foci? Is a single CRE enough? Does the presence of multiple CREs increase transcriptional output? Answering these questions will likely require the ability to localize both trans-factors and multiple CREs simultaneously with high spatial resolutions, and the use of perturbation methods (e.g. optogenetic manipulation of the low-complexity domains that mediate protein-protein interactions (Shin et al. 2017, 2018)). The next section describes the mechanisms that may be involved in the formation of transcription-associated nuclear foci and CRE-hubs.

Phase separation and nuclear structure

Phase separation has long been known as a process of self-organization in cells, and can explain the formation of membraneless compartments in the nucleus. Some of the most prominent examples include nuclear bodies (e.g. nucleoli, Cajal bodies and DNA damage repair sites (Hyman et al. 2014; Banani et al. 2017; Boeynaems et al. 2018; Shin and Brangwynne 2017)) as well as chromatin itself (Gibson et al. 2019). Phase separation offers unique opportunities for controlling biochemical micro-environments by locally increasing the concentration of the constituents (and consequently chemical reaction rates), while still permitting dynamic exchange of reactants and products. Different mechanisms, involving distinct types of molecular interactions, have been proposed to lead to phase separation in the nucleus (Banani et al. 2017; Erdel and Rippe 2018; Kato and McKnight 2017).

Lately, liquid-liquid phase separation (LLPS) has received much attention as a possible mechanism of nuclear organization. LLPS compartments typically have spherical shapes, can fuse together, can deform under shear flow, are in exchange with the surrounding medium, and show a dynamic internal organization (Hyman et al. 2014). The key molecular driving force for the formation of liquid-like condensates are weak, multivalent interactions, typically involving proteins with intrinsically-disordered domains (IDRs) (Shin and Brangwynne 2017; Banani et al. 2017).

Spatial clustering of super-enhancers often relies on the formation of condensates displaying typical properties of LLPS (Chong et al. 2018; Boija et al. 2018; Cho et al. 2018; Sabari et al. 2018; Hnisz et al. 2017). Consistently, the proteins involved in the formation of these condensates (TFs and coactivators) often contained IDRs that facilitated their nucleation. In this scenario, activating/repressive signals from CREs within a condensate may be transmitted over relatively large distances to the transcription machinery without requiring direct physical interactions. So far, it is unclear on which timescale CREs remain associated with protein foci or whether promoters only need brief and transient encounters with these
condensates for transcriptional activation (Cho et al. 2018). A more thorough discussion of the roles of LLPS in transcription can be found elsewhere (Mir et al. 2019).

A second mechanism that can drive the formation of transcription-associated foci and CRE-hubs is polymer-polymer phase separation (PPPS) (Erdel and Rippe 2018). This mechanism is based on the binding of soluble ‘bridging factors’ to a long polymer chain, connecting two or more sites on the polymer. If the density of cross-links becomes sufficiently high, a polymer collapse takes place, resulting in a locally compact polymer globule. In this model, transient binding of cross-linking factors and their constant exchange with the surrounding nucleoplasm still leads to stable condensates, as long as a steady state with a sufficiently high density of bridging interactions is maintained (Erdel and Rippe 2018). Recently proposed models of phase separation for transcriptional control (Hnisz et al. 2017) are compatible with PPPS as they do not set any requirements on the chemical nature of the crosslinks between chromatin chains. Counterintuitively, factors able to assemble liquid droplets in vitro may still form PPPS in vivo (Erdel et al. 2020). Despite their association to DNA, PPPS condensates are highly dynamic and, in some cases, can require the energy of ATP hydrolysis to ensure proper sub-nuclear positioning (Guilhas et al. 2020).

Perhaps the most important difference between LLPS and PPPS is that condensates formed via LLPS can persist independently of the chromatin scaffold, while condensates formed via PPPS strictly rely on such a scaffold and would disassemble in its absence (Fig. 3C). Future studies should experimentally test LLPS and PPPS models by examining concentration dependency and the need for a chromatin scaffold (Guilhas et al. 2020; Erdel et al. 2020; Erdel and Rippe 2018). These experiments would be particularly timely to further clarify whether the formation of CRE-hubs requires transcription-associated foci and vice-versa.

3. Time scales of nuclear dynamics

With the advent of high-resolution live imaging technologies, measurements of dynamic nuclear phenomena involving chromatin and binding factors not only have become a possibility but already a reality. These involve, on the one hand, measurements of, e.g., transition rates, binding/unbinding of factors, and more generally of kinetic rate constants that provide kinetic but not spatial information (Senecal et al. 2014; Zoller et al. 2018). On the other hand, it is now possible to go a step further to get real-time dynamic measurements with spatial information that, e.g. follow the movement of chromatin or molecules in space and time (Shaban et al. 2020; McCord et al. 2020). Here, we distinguish between cis-dynamics, that happen on the same chromosome, such as enhancer–promoter interactions or chromatin loop-formation, and trans-dynamics that involve interactions of trans-acting factors with chromatin, the formation of higher-order structures such as CRE-hubs, or transvection.
Several studies, using both high-resolution and live-cell imaging, but also single-cell 3C-based methods, have revealed a highly heterogeneous nature of genome order in both space and time (Cattoni et al. 2017; Flyamer et al. 2017; Nagano et al. 2013; Stevens et al. 2017; Bintu et al. 2018; Hansen et al. 2018; Finn et al. 2019). However, few studies have in fact succeeded in quantifying actual chromosome dynamics in single cells, leaving a multitude of unanswered questions that are crucial for understanding genome organization and function. For example, the timescales over which the structure of chromosomes change remain largely unknown. How are single cell chromosome topologies established during the cell cycle and during development? How does loop interaction frequency change across different timescales? How does the dynamic organization of the genome relate to gene expression and other nuclear processes, such as cell-type specification, DNA repair or replication?

To directly address these and other questions, a number of imaging and labeling techniques have been put forward (Fig. 4). Following chromatin dynamics in space and time not only requires development of state-of-the-art imaging technologies that often go beyond the diffraction limit (Lakadamyali and Cosma 2020; Brandão et al. 2020), but also development of highly sophisticated chromatin labeling capabilities that typically require a strenuous and time-consuming combination of molecular cloning, genome editing and genetics (Sato et al. 2020; Shaban and Seeber 2020).

Several methods have been developed to visualize global chromatin dynamics at the nuclear scale (Zidovska 2020; Shaban et al. 2020). Imaging the dynamics of multiple genomic sites requires tools to fluorescently label specific sequences. Early methods for live imaging of
specific genomic loci include the endogenous insertion of large binding site arrays for fluorescently tagged Lacl or TetR repressors (Marshall et al. 1997; Heun et al. 2001; Chubb et al. 2002; Chuang et al. 2006; Kumaran and Spector 2008; Masui et al. 2011). Other approaches that require genome editing use the ParB/parS or ANCHOR DNA-labeling systems (Saad et al. 2014; Germier et al. 2017; Chen et al. 2018). In addition, catalytically inactive Cas9 enzymes (dCas9) tagged with GFP have been used to target specific genomic loci in living cells (Chen et al. 2013; Stanyte et al. 2018; Gu et al. 2018; Ma et al. 2018). These efforts led to the first direct measurements of CRE dynamics in living cells (Lucas et al. 2014; Germier et al. 2017; Herbert et al. 2017; Chen et al. 2018; Gu et al. 2018; Alexander et al. 2019; Khanna et al. 2019; Li et al. 2020a; Lim et al. 2018), assessments of large-scale chromatin dynamics (Zidovska et al. 2013; Nozaki et al. 2017; Shaban et al. 2018; Zidovska 2020) and quantification of DNA-binding factor interaction dynamics (Lionnet and Wu 2021) reviewed by (Lionnet and Wu 2021).

Many improvements in both microscopy and probe development are linked by their ability to control the photon budget (Planchon et al. 2011; Zhao et al. 2011; Lavis 2017): the number of detectable photons that a particular fluorophore contributes to the experiment, which is limited because of photo-chemistry and photobleaching (i.e. the permanent loss of fluorescence due to photo-induced chemical changes). However, further increases in localization precision, imaging rate, and imaging time, all require the detection of higher numbers of photons. This generates an optimization dilemma between resolution, speed, depth of view, and photodamage (Fig. 5A). Thus the limiting factor is not the microscope; it is the photon budget. Close attention needs to be paid to the fluorophore choice because the photon budget is one of the most commonly neglected factors that would significantly affect the feasibility and success of an experiment. A common strategy to sidestep this problem is to increase the number of fluorophores bound to individual molecules.

**Cis-dynamics**

The first pioneering work analyzing the dynamics of chromosomal loci was performed in lymphocyte B cells, where either V_{H} or D_{H}J_{H} regions at the immunoglobulin gene loci were followed one-by-one by single-particle tracking (Lucas et al. 2014). These loci were shown to display a subdiffusive behavior with fractional Langevin motion and were mostly spatially confined. Similarly, in live mouse embryonic stem cells the Fgf5 enhancer and promoter displayed subdiffusive behavior and their motility increased during differentiation to epiblast-like cells concomitant with transcriptional activation (Gu et al. 2018). The authors proposed that higher diffusivity of CREs increases stochastic encounters within TADs, potentially boosting successful enhancer–promoter interactions.

To partially avoid biases introduced by the inherent large-scale three-dimensional motion of the entire nucleus, several efforts used two-color imaging to monitor the relative motion of two genomic loci on the same chromosome (Figure 4C). This dual labeling strategy was used in lymphocyte B cells to characterize the relative movement of V_{H} and D_{H}J_{H} regions (Khanna et al. 2019). V_{H} and D_{H}J_{H} elements spatial positions fluctuate locally while their distance remains nearly constant over time, but abrupt changes in motion could be observed and suspected to originate from rapid temporal changes in large-scale chromatin conformations. These studies indicate that chromatin dynamics is largely subdiffusive in mammalian cells, with occasional abrupt changes in motion.

In mouse embryonic stem cells, live two-color imaging revealed that the Pou5f1 and Sox2 enhancers are frequently in proximity with their target gene transcription site (100–200 nm) (Li et al. 2020a). The authors argued that this 100–200 nm-sized cluster of enhancers could
concentrate components of the transcription machinery and activate gene expression (Li et al. 2020a). Alexander et al. monitored the dynamics of the Sox2 gene in mouse embryonic stem cells and its essential endogenous SCR enhancer, positioned ~100 kb away to show that the Sox2 enhancer–promoter spatial organization exhibited a high cell-to-cell variability (Alexander et al. 2019). The two loci displayed sporadic sharp topological transitions within a tightly confined space in the ~200 nm length scale. These observations suggest that enhancer–promoter pairs could be frequently and transiently interacting with each other, in contrast with the classical stable contact model (Deng et al. 2012, 2014; Bartman et al. 2016). However, it is unclear if loop extrusion and chromatin conformation dynamics impact these transient enhancer–promoter interactions and what are the consequences on transcription.

**Trans-Dynamics**

Another approach to decipher the time-dependent nature of the genome is to quantify the kinetics and dynamics of trans-acting factors, such as those required for loop extrusion, TAD formation, or stabilization of enhancer–promoter interactions. Their dynamics could potentially offer a route to study these processes and thus shed light on the underlying chromatin mechanics. For example, using single-molecule imaging approaches in live cells, cohesin’s average residence times (~20 min) are around an order of magnitude longer than those of CTCF (~1 min) (Hansen et al. 2017). These residence times are relatively stable compared to conventional transcription factors that can bind and dissociate from DNA on timescales of seconds (Liu et al. 2014; Mazza et al. 2012; Hansen et al. 2018; Chen et al. 2014) but highly dynamic compared to the length of the cell cycle (~24 hours), suggesting that chromatin loops dynamically form and break multiple times throughout the cell cycle (Hansen et al. 2018; Zhang et al. 2019). In addition, loop extrusion by the cohesin complex occurs at a maximum rate of ~2 kb/sec *in vitro* (Davidson et al. 2019; Kim et al. 2020), thus loops could likely play a role in the dynamics of enhancer–promoter interactions.

Several studies showed that components of the transcription machinery, such as RNA polymerase II, the Mediator complex, BRD4, and transcription factors form transcription-associated nuclear phase-separated condensates (Chong et al. 2018; Sabari et al. 2018; Cho et al. 2018) (see also Section 2 above). Using live-cell single molecule imaging, Chong et al. showed that interactions between low-complexity regions of TFs are highly dynamic, typically on the second to minutes time scale. Consistently, Cho et al. used single-particle tracking to show that Mediator and RNA PolII form large (> 300 nm), long-lived (> 100 s), chromatin-associated clusters where they co-localize in a transcription-dependent manner. Mediator clusters display properties of phase-separated condensates and may mediate transient enhancer–promoter communication over large distances (few hundred nanometers) (Fig. 3). Critically, in some instances, transcription-associated foci are highly dynamic and can form and dissociate in a matter of seconds (Liu et al. 2014; Cho et al. 2018; Dufourt et al. 2018). How these short-lived foci contribute to establishing long-range interactions between CREs remains an open question.

**4. Functional consequences of spatiotemporal chromosome organization**

After static and dynamic descriptions of nuclear organization, the next frontier is to provide a connection to biological function. What is the functional impact of nuclear order? Is there an evolved functional relationship between chromatin dynamics and transcription? To address these and similar questions we need the ability to assess physical and genomic structure, chromatin dynamics and spatial distributions of regulatory factors, as well as biological function—ideally all simultaneously—which is a highly challenging task. Most likely, and
similar to the dilemma of sharing a photon budget for imaging experiments, designing one experiment that can accomplish all these tasks simultaneously still lies somewhere in the future, and with current technologies one has to make choices about which of these features to optimize (Figure 5B). Here we focus on initial progress to assess biological function in terms of measurements of transcription and its relationship to local genome organization.

![Figure 5. (A) The photon budget creates a tug-of-war between different desirable optimization strategies in photonic imaging. While the ideal imaging experiment would push the limits of each property on the corners of the tetrahedron, the finite size of the photon budget sets a fundamental limit on the combined optimization of all four corners. Pushing one corner to its extreme limit implies giving up on the performance of the others. (B) The (epi)-genome contains all the information necessary to produce an entire organism. However, characterization of the functional output of the genome requires the ability to monitor multiple observables simultaneously. These include nuclear architecture, dynamics, occupancy by regulatory factors, and the biological output. Current technologies (showing here only a small number for simplicity) hardly accomplish a subset of these tasks in a single experiment.](image)

Arguably one of the most functionally-relevant outputs of the nucleus is its genetic program, which is executed via gene activity. Transcriptional dynamics can be monitored by in vivo visualization of nascent mRNA transcripts using bacteriophage-based reporter cassettes, as pioneered over twenty years ago (Bertrand et al. 1998; for a review see: Sato et al. 2020). MS2 and PP7 stem-loops are positioned in the gene body and coexpressed with a fluorescently-tagged coat protein to visualize nascent transcription in live cells (Janicki et al. 2004; Larson et al. 2011; Bothma et al. 2014). Measuring the signal intensity of the fluorescent foci at the site of gene activity was used to show that transcription is a stochastic process, subject to molecular noise, that exhibits transcriptional bursts with intervals of mRNA production followed by intervals of transcriptional inactivity (see review (Rodriguez and Larson 2020)). This stochasticity results in gene expression noise and cell-to-cell variability, the causes of which could possibly involve spatiotemporal fluctuations of the chromosome (Shah et al. 2018) or of CRE-hubs (Shah et al. 2018).

A powerful application of this approach showed that enhancers control the frequency of transcriptional bursts, are able to coactivate linked genes, and yet can exhibit large spatial separation from their target genes even during transcriptional activation (Fukaya et al. 2016). These results contribute to the conflicting picture about the role of chromatin topology on genome function (Lupíañez et al. 2015; Northcott et al. 2014; Franke et al. 2016; Nora et al. 2017; Rao et al. 2017; Schwarzer et al. 2017). While it is clear that enhancers and promoters can come into spatial proximity, and that physical proximity is somehow linked to transcriptional activation, such evidence is often correlative. Is enhancer–promoter proximity a consequence of transcriptional activation or is it needed for transcriptional activation? And if
so, does transcription happen simultaneously with proximity or is it uncoupled? Even if CRE proximity is necessary for transcriptional activation, is it also necessary for sustained activity?

Answering these questions will most likely require simultaneous detection of transcription and genome organization in the same cell. This is particularly important for enhancer–promoter interactions and the underlying mechanisms governing transcriptional control. Recent development of elaborate imaging-based methods – either via live-cell imaging or high-resolution fixed-tissue localization-based microscopy – have enabled the first direct visualization of long-range enhancer–promoter interactions coupled with transcriptional activity (Chen et al. 2018; Alexander et al. 2019; Mateo et al. 2019; Espinola et al. 2021; Li et al. 2020a; Barinov et al. 2020). Hi-M and ORCA were used to visualize both chromatin structure and transcriptional activation in Drosophila to show that transcriptionally active and inactive cells display very similar enhancer–promoter proximity (Mateo et al. 2019; Espinola et al. 2021). While these approaches enable the detection of transcription and the topology of multiple enhancers and promoters, they do not shed light on their dynamics.

To observe whether two distal chromosomol regions interact in a functionally significant manner, the notions of proximity or contact may no longer be sufficient. Rather, a more complex imaging assay is needed that involves simultaneous live image capture of three differently colored DNA tags: two to dynamically follow the motion of the distal chromosomal sites, such as an enhancer and a promoter, with the third tag serving as a reporter for functional proximity, i.e. it only lights up when specific events, such as transcription, occur. Here the MS2/PP7 labeling systems are key in two ways: analysis of the fluorescence intensity signal for function and simultaneously tracking its location to gain insights into spatial relationships.

Using such an approach, progress towards a causal connection between dynamic enhancer–promoter communication and gene expression has recently been carried out. Live imaging experiments in Drosophila embryos visualized physical enhancer–promoter interactions and transcription at the eve locus in Drosophila embryos (Chen et al. 2018). Sustained physical proximity between the enhancer and the promoter of a distal reporter gene was shown to be necessary for transcriptional activity of the reporter gene. In addition, transcriptional activity also seems to stabilize this proximal conformation and have an impact on the physical size of the active gene locus, suggestive of a reciprocal interplay between enhancer–promoter dynamics and transcriptional activity (van Steensel and Furlong 2019).

In contrast, a similar study in mouse embryonic stem cells reported that enhancer–promoter proximity at the Sox2 locus is uncoupled from transcription (Alexander et al. 2019). In line with these results, 3D DNA-FISH and chromosome conformation capture in fixed cells revealed a decreased spatial proximity between the Shh gene and its enhancers during the differentiation of mouse embryonic stem cells into neural progenitor cells (Benabdallah et al. 2019). It would be interesting to analyze enhancer–promoter interactions at the Shh locus using live imaging to determine the dynamics of these interactions at timescales shorter than days of cell differentiation, to see if they are transient and unstable, as for the Sox2 locus. In many loci several enhancers can activate transcription of a promoter, individually or conjointly. Thus an observed lack of proximity may not be a good indicator for the need of enhancer–promoter proximity as some other enhancer in the neighborhood or within a CRE-hub may be activating transcription.

It is still an open question what exactly contact or proximity mean in the context of an active enhancer–promoter pair, and how these notions relate to the emerging evidence of
transcription-related CRE-hubs and transcription-associated protein foci. Again, to discern between these models, and to devise new ones, dynamic measurements would need to simultaneously track multiple CREs, and factors in space and time. Testing current models of enhancer–promoter communication will also require imaging methods with high-resolution to distinguish enhancer–promoter loops in physical interaction range (in and of itself an ill-defined length scale) from observed enhancer–promoter proximity (150–200 nm). Further progress towards this issue has been put forward using multi-color localization microscopy to achieve 1–2 kb resolution in an 18 kb gene locus in *Drosophila*, where the transcriptionally active enhancer–promoter pair was still more than 150 nm separated, thus again not in physical interaction range (Barinov et al. 2020).

At the time of writing this chapter, there are still a number of unanswered questions, in particular, if the observed spatial gaps between active enhancers and promoters prove to be a general feature of active transcription-associated loci. In that case, the most pressing question is to understand how information about activity spreads across large spatial distances (>100 nm) from the transcription factor-bound active enhancer to the transcription-engaging promoter. It may take another generation of experiments and maybe even technologies to make progress in the understanding of how the general architecture of the folded nuclear genome regulates genome function and vice versa. The observations above could suggest that gene regulation by stable enhancer–promoter looping is not a generality, and more elaborate models of long-range enhancer–promoter communication should take other spatial considerations or constraints into account.

**Concluding remarks**

The last two decades have seen a revolution in optical imaging approaches, with the advent of super-resolution microscopy, live tracking of individual molecules, and multiplexed methods. Application of these approaches already made a strong impact on our understanding of the interplay between the structure and dynamics of the nucleus and its transcriptional output. However, understanding causality and gaining an integrated picture of how dynamic changes in chromosome organization control the timing and levels of transcription will require considerable further advances in several interdisciplinary areas. A crucial issue will be the necessity to measure in individual cells the multiple critical components of nuclear organization underlying gene activity: transcriptional output, spatial arrangement of CREs, dynamic long-range chromosomal interactions, and the epigenetic state of chromatin (Figure 5B). Many of these quantities can already be measured in single cells, one-by-one or in pairs, but detecting them simultaneously presents a long-term challenge that will require significant advances in design and development of microscopy, as well as in labeling and genome editing tools.

One area that is still largely underdeveloped is the single-cell study of the dynamics and genome occupancy of trans-regulatory factors and how they influence genome organization and function. For example, it will be crucial to monitor the dynamic binding of multiple transcription factors to CREs to understand how the latter turn into a state of activity that ultimately leads to a transcriptional output. Likewise, it will be critical to further integrate how signalling, environmental cues, or splicing shape the assembly and dynamics of CRE-hubs and transcription associated foci to regulate transcription.

There are still a lot of open questions and many new frontiers to explore. Progress in optical and electron microscopies have already achieved the spatial resolutions necessary to start revealing the organization of chromatin in cells with a resolution of a few nanometers (Ou et
al. 2017; Jungmann et al. 2016; Ricci et al. 2015). However, visualization of chromatin structure at the nanometer-scale as inert/“feature-less” beads may not be enough: to get a full understanding of the role of chromatin structure for biological function will require the detection of multiple molecules (e.g. proteins, RNA, DNA) with sequence specificity, nanometer precision, and dynamic probes. These developments have to go hand in hand with improvements in labeling and sample preparation procedures that conserve structures at the molecular scale. In addition, still largely lacking are perturbation techniques that can directly probe causal relationships between nuclear structure, the environment and function. For example optogenetic association or dissociation of molecular compounds or DNA loci should provide powerful handles for causality experiments. Excitingly, all these advances should bring a deeply needed facet to tracking single molecules interacting with DNA and asserting control over transcription programs, one factor at a time.

**Acknowledgements**

This work was supported in part by the European Union’s Horizon 2020 Research and Innovation Program (Grant ID 724429) (M.N.), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, project ID 431471305) (M.G.), the Laboratoire d'Excellence Revive (Investissement d'Avenir; ANR-10-LABX-73) (I.B.), the US National Science Foundation, through the Center for the Physics of Biological Function (PHY–1734030) and by National Institutes of Health Grants R01GM097275, U01DA047730, and U01DK127429 (TG).
References


Neems DS, Garza-Gongora AG, Smith ED, Kosak ST. 2016. Topologically associated domains enriched for lineage-specific genes reveal expression-dependent nuclear


Rowley MJ, Corces VG. 2016. The three-dimensional genome: principles and roles of


http://dx.doi.org/10.1126/science.aar3958.


