- -

The three-dimensional genome: regulating gene expression during pluripotency and development

- -

Andrey, Guillaume; Mundlos, Stefan

REVIEW

The three-dimensional genome: regulating gene expression during pluripotency and development

Guillaume Andrey[1] and Stefan Mundlos[1,2,*]

ABSTRACT

The precise expression of genes in time and space during embryogenesis is largely influenced by communication between enhancers and promoters, which is propagated and governed by the physical proximity of these elements in the nucleus. Here, we review how chromatin domains organize the genome by guiding enhancers to their target genes thereby preventing non-specific interactions with other neighboring regions. We also discuss the dynamics of chromatin interactions between enhancers and promoters, as well as the consequent changes in gene expression, that occur in pluripotent cells and during development. Finally, we evaluate how genomic changes such as deletions, inversions and duplications affect 3D chromatin configuration overall and lead to ectopic enhancer-promoter contacts, and thus gene misexpression, which can contribute to abnormal development and disease.

KEY WORDS: 3D chromatin structure, Enhancers, Gene regulation, Non-coding mutations, Stem cells

Introduction

The expression of developmental genes is tightly orchestrated in time and space throughout embryonic development as well as organogenesis, and transcriptional regulation is a well-recognized essential component of normal embryogenesis. Complex changes in expression are associated with, and are essential for, determining cell fate, lineage commitment, the establishment of the body plan, and organ formation (Zeller et al., 2009; Petit et al., 2017). These dynamic transcriptional processes are mainly mediated by promoters that are located near the transcription start site (TSS) and by regulatory regions termed enhancers that can be located at a considerable distance from the TSS. Once activated by specific transcription factors, enhancers produce spatially and temporally restricted patterns of activity that determine the expression of their target genes (Spitz and Furlong, 2012). They regulate gene expression by physically interacting with their target gene promoters via looping of DNA. This looping and the associated physical interaction between enhancers and promoters is propagated by proteins that facilitate this contact, such as the mediator complex (see Glossary, Box 1) (Carlsten et al., 2013). Importantly, changes in these regulatory units are thought to play a role in the evolutionary development of new or altered patterns of gene expression and the subsequent acquisition of morphological novelties (Chan et al., 2010; Long et al., 2016).

Together, promoters and enhancers constitute the regulatory landscapes of genes. At developmentally active sites, these landscapes can span several hundreds of kilobases and can contain numerous enhancers (Montavon et al., 2011; Andrey et al., 2013; Marinić et al., 2013; Franke et al., 2016; Symmons et al., 2016). Such loci require an exquisitely complex regulation of their 3D structure, to assure precise and robust gene expression. In recent years, chromosome conformation capture technologies (e.g. 4C, HiC and its derivatives) have been used to quantify the frequency of chromatin contacts at specific loci or genome wide (Sati and Cavalli, 2017). These studies have shown that the genome is organized in a non-random fashion, resulting in a higher order 3D organization of entire chromosomes. On a sub-chromosomal level, chromatin domains exhibiting high levels of internal interactions have been identified using a variety of algorithms and definitions. One particular type of chromatin domain, referred to as a topologically associating domain (TAD, see Glossary, Box 1), has emerged as being particularly important. TADs are defined as chromatin domains with high internal interactions and are separated from each other by regions of low interaction called boundary elements. TADs have been shown to be a fundamental component of the 3D organization of the genome in the nuclear space, restraining and facilitating enhancer-promoter interactions mostly in a cell type-independent manner (Dixon et al., 2012; Nora et al., 2012). During embryonic development, in particular upon cell differentiation and lineage commitment, the micro-architecture of TADs becomes more structured, probably because TADs become directly involved in transcriptional control by directing enhancer-promoter contacts (Bonev and Cavalli, 2016).

Characterization of the various components that control gene regulation, i.e. enhancers, promoters and their associated 3D interactions, has become a standard approach to deciphering the mechanisms underlying transcriptional control and has thus provided new frameworks for understanding basic developmental processes, disease pathology and the evolution of organs and structures (Montavon et al., 2011; Andrey et al., 2013; Lonfat et al., 2014; Lupiáñez et al., 2015; Franke et al., 2016). Here, we review the general mechanisms of genome organization and gene regulation during vertebrate development and cell differentiation. We describe the function and organization of enhancers and discuss current concepts of 3D chromatin dynamics during development and lineage commitment, focusing on TADs and their involvement. Finally, we review the different types of mutations that can affect the integrity of chromatin structure and thus result in gene misexpression, developmental abnormalities and disease.

Enhancers and promoters form regulatory landscapes

The regulation of developmental or lineage-specific genes is tightly controlled through the activities of regulatory elements called enhancers. Enhancers were originally described as genomic regions that can drive the transcription of a reporter gene in specific reporter constructs (Banerji et al., 1981). Such a construct would classically

[1]Max Planck Institute for Molecular Genetics, RG Development & Disease, 14195 Berlin, Germany. [2]Institute for Medical and Human Genetics, Charité Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany.

*Author for correspondence (mundlos@molgen.mpg.de)

G.A., 0000-0002-0911-4907; S.M., 0000-0002-9788-3166

DEVELOPMENT

Box 1. Glossary

A/B compartments. Chromosomes are organized into active (A) and inactive (B) compartments that interact in a homotypic fashion (Lieberman-Aiden et al., 2009).

ATAC-seq. Assay for transposase accessible chromatin with high-throughput sequencing (ATAC-seq) that uses *in vitro* transposition of adaptors into open chromatin regions, followed by deep sequencing to map open regions at a genome-wide level (Buenrostro et al., 2013).

ChIP-seq. An approach that combines chromatin immunoprecipitation of the desired epitope with deep sequencing to identify genomic regions bound by a specific transcription factor or marked by a specific histone post-translational modification.

Chromocenters. Clusters of pericentromeric satellite repeats from different chromosomes that are found in specific nuclear landmark structures.

Cohesin complex. A ring-like multiprotein complex involved in the pairing of sister chromatid during cell divisions and in the production of chromatin loop between distal DNA segments.

CTCF. An 11-zinc-finger transcription factor that is enriched at boundary elements of TADs and loop domains.

DNA replication time zones. Regions of the genome that are replicated at distinct times during S phase.

DNAse-seq. A method that captures nucleosome-depleted fragments digested by the DNAse I enzyme and uses deep sequencing to map chromatin open regions at a genome-wide level (Song and Crawford, 2010).

Embryonic stem cells (ESCs). Pluripotent stem cells that are derived from the inner cell mass of a blastocyst and can differentiate into any cell type.

Induced pluripotent stem cells (iPSCs). Pluripotent stem cells that are generated by the reprogramming of differentiated cells via the introduction of several pluripotency factors.

Insulated neighborhoods. Chromatin domains defined by chromatin loops formed by a CTCF homodimer co-bound by cohesin and containing at least a gene (Hnisz et al., 2016).

Lamina-associated domains (LADs). Chromatin domains that interact with the lamina of the nucleus.

Loop domains. Domains of preferential interaction that are marked by a loop at their corner (Rao et al., 2014).

Mediator complex. A large protein complex with a variable subunit composition, which controls various processes necessary for transcription, including 3D chromatin architecture.

Polycomb or Polycomb group (PcG) proteins. Proteins that can remodel chromatin so that epigenetic repression can occur. They play key roles in stem cell identity and cell differentiation (Di Croce and Helin, 2013).

Polycomb repressive complex 1 (PRC1) and 2 (PRC2). PRC1 contributes to chromatin compaction and catalyzes the monoubiquitylation of histone H2A. PRC2 is recruited to chromatin, contributes to chromatin compaction and catalyzes the methylation of H3K27 (Margueron and Reinberg, 2011).

Topologically associating domains (TADs). Chromatin domains that exhibit high levels of internal interactions. They are separated from each other by regions of low interaction termed boundary elements (Nora et al., 2012).

consist of the DNA fragment to be tested, a minimal promoter (often from the β-globin gene), and a reporter gene, such as *lacZ* or *GFP* (Shlyueva et al., 2014). Once introduced into the genome of an embryo, the expression of this reporter construct is expected to reflect the regulatory activities of the tested region. Large-scale studies using this approach have identified thousands of enhancers that drive tissue- and time-specific gene expression patterns in mice or flies (Visel et al., 2007; Manning et al., 2012). Techniques that assess chromatin openness or post-translational histone modifications can also be used to localize putative enhancers (see Box 2). Indeed, as transcription factors preferentially bind to open

chromatin, techniques that detect these regions, such as ATAC-seq or DNAse-seq (see Glossary, Box 1), are very useful for establishing genome-wide enhancer footprints (Song and Crawford, 2010; Buenrostro et al., 2013). Another approach utilizes the acetylation of histone 3 at lysine 27 or 122 (H3K27ac or H3K122ac), which correlates with active enhancer activities and can be probed by chromatin immunoprecipitation followed by deep sequencing (ChIP-seq; see Glossary, Box 1) (Creyghton et al., 2010; Rada-Iglesias et al., 2011; Pradeepa et al., 2016).

Enhancers are bound by tissue-specific transcription factors and can thereby produce highly controlled regulatory outputs, in time, in space or in specific cell types (Spitz and Furlong, 2012). The birth of new enhancer regions, through the acquisition of transcription factor binding and subsequent novel regulator function, is thus thought to be a key component in gene neo-functionalization and a driver for evolution (Long et al., 2016). According to current concepts, complex expression patterns of developmental genes are achieved by the additive effect of multiple enhancers, with each element regulating a subset of the global expression pattern of a gene (Fig. 1). For example, *in vivo* studies of the α-globin gene locus, which is controlled by a cluster of enhancer elements, also referred to as a super enhancer, have shown that each enhancer seems to act independently and in an additive fashion without clear evidence of synergistic or higher-order effects (Hay et al., 2016). A recent comprehensive study of the Indian hedgehog (*Ihh*) locus also shows that *Ihh* is regulated by a multipartite enhancer ensemble consisting of at least nine enhancers with individual combinations of tissue specificities that function in an additive manner (Will et al., 2017). Furthermore, this study showed that enhancers function in a dosage-dependent manner; an increase in the copy number of enhancers results in increased gene expression, although this increase is not linear and differs in different tissues. Likewise, a significant loss of transcription of the *PIM1* leukemic oncogene in a human cell line was only obtained through the combinatorial repression of several weak enhancers embedded within a super-enhancer region and not by the repression of single enhancers (Xie et al., 2017). Enhancer crosstalk is also sometimes required to control gene transcription properly. At the mouse *Krox20* (also known as *Egr2*) locus, for example, an enhancer element is required to potentiate another enhancer that acts in an auto-regulatory fashion to sustain *Krox20* expression (Thierion et al., 2017). In addition, strong enhancers at the *Drosophila knirps* and *hunchback* loci were shown to act in a sub-additive manner by buffering each other's activities thereby allowing a constant transcriptional output (Bothma et al., 2015). Thus, the regulatory landscapes of developmental genes seem to consist of complex enhancer ensembles that act in an interactive manner to achieve the precise and robust control of gene expression that is essential for normal development.

In vertebrates, enhancers often locate at a distance from the promoter they control and bypass nearby genes (de Laat and Duboule, 2013). Indeed, a genome-wide study in human cell lines has estimated that as few as 7% of distal regulatory elements control their closest promoters (Sanyal et al., 2012). The specificity of promoter elements for certain enhancers might thus play a role in directing their regulatory activities. A study of the *HoxD* cluster during mouse limb development, for example, determined that each of the *HoxD* gene promoters, despite being very similar, has a different affinity for the same set of digit enhancers (Montavon et al., 2008). Furthermore, pioneer studies in *Drosophila* have shown that promoters can each respond in a different fashion to particular types of enhancer (Zabidi et al., 2015; Arnold et al.,

Box 2. Approaches and techniques for enhancer discovery

Enhancers feature several hallmarks that allow their identification:

Open chromatin. Enhancers are found at sites of open chromatin, which can be identified using approaches such as DNAse-seq or ATAC-seq.

Chromatin modifications. Poised and active enhancers are enriched for the chromatin modification H3K4me1. Active enhancers are further modified with H3K27ac and/or H3K122ac (Heintzman et al., 2009; Creyghton et al., 2010; Rada-Iglesias et al., 2011; Pradeepa et al., 2016). These modifications can be mapped in the appropriate cell type using ChIP-seq.

Transcription factor binding. Enhancers are bound by transcription factors, which tether chromatin remodeling and transcriptional complexes (Spitz and Furlong, 2012). ChIP-seq for transcription factors that are involved in specific biological process can be used to determine the genomic location of enhancers (Araya et al., 2014).

Conservation. Because of the evolutionary constraints imposed on enhancer sequences by the requirement to bind transcription factors, enhancers are predicted to be more conserved genomic regions (Pennacchio et al., 2006).

Other techniques and approaches that can be used to identify and characterize enhancers include:

Reporter assays. In this approach, putative enhancer sequences are placed in front of a minimal promoter and a reporter gene. By integrating such a construct into the genome, it is possible to detect, in a developing embryo, the tissue-, cell- or time-specific activity of the reporter gene, which mirrors the activity of the enhancer.

STARR-seq. In this approach, which takes advantage of the fact that enhancers can work independently of their position and orientation, the candidate enhancer region is placed downstream of the promoter. Accordingly, genome-wide libraries of fragments can be cloned and active enhancer regions that can induce their own transcription can thus be measured by the abundance of transcripts (Arnold et al., 2013).

CAGE. Cap analysis of gene expression is a method that can detect enhancer activities by measuring the presence of bidirectional capped RNA transcripts (Andersson et al., 2014).

2017). Accordingly, the distance between enhancer and promoter is not the primary parameter that controls gene activation. A well-characterized case of long-range regulation occurs at the mouse sonic hedgehog (*Shh*) locus, where the limb-specific enhancer of *Shh* (called ZRS) is located more than 1 Mb away from the *Shh* promoter (Lettice et al., 2003; Sagai et al., 2005). Similarly, at the mouse *HoxD* cluster, a regulatory landscape that is distributed across a 750 kb region controls the digit-specific expression of the *Hoxd13* to *Hoxd10* genes (Montavon et al., 2011), and at the mouse *Sox9* locus, a 1.7 Mb region controls the intricate expression pattern of this gene (Franke et al., 2016).

The insertion of regulatory sensors, which consist of a minimal promoter and a reporter gene, at multiple sites in a given genomic locus has been used to determine the extent of such regulatory landscapes (Kondo and Duboule, 1999; Ruf et al., 2011; Akhtar et al., 2013; Marinić et al., 2013; Symmons et al., 2016). In these experiments, one can observe that the regulatory inputs from enhancers are sensed over hundreds of kilobases with varying intensities. These and other studies have also revealed that regulatory boundaries exist, as demonstrated by a decrease and ultimately the disappearance of the reporter signal. The regions of activity that exist between such boundaries have subsequently been shown to correspond to TADs (discussed in detail below) (Dixon et al., 2012; Nora et al., 2012; Tsujimura et al., 2015; Lupiáñez et al., 2016). In conclusion, regulatory landscapes can extend over hundreds of kilobases, are part of or defined by the boundaries of a TAD, and characteristically consist of numerous regulatory

elements that together control the expression of a target gene. These landscapes have also been referred to as regulatory archipelagos or holo enhancers (Montavon et al., 2011; Marinić et al., 2013; Hughes et al., 2014) (Fig. 1).

Chromatin domains as basic genomic regulatory units

In recent years, various technological approaches have shown that the genome is organized in a non-random fashion into higher order chromatin domains. For instance, proximity ligation technologies (see Box 3) have been used to determine the frequency at which two genomic regions can be crosslinked and are thus in close proximity to each other in the nucleus (Sati and Cavalli, 2017). Specifically, HiC and other C-technologies have revealed the existence of self-associating and insulated chromatin domains called TADs, which are separated by regions of low interactions called boundaries (Dixon et al., 2012; Nora et al., 2012). These Mb-size structures associate with multiple genomic features, such as functional chromatin modifications, DNA replication time zones, chromocenters, or lamina-associated domains (LADs; see Glossary, Box 1) (Dixon et al., 2012; Pope et al., 2014; Wijchers et al., 2015). The existence of TADs has also been confirmed by another approach that does not rely on chromatin crosslinking and ligation but on DNA sequencing from a collection of thin nuclear sections (Beagrie et al., 2017). Strikingly, TADs remain largely unchanged during lineage commitment, are stable across different cell lines, and are even conserved between species (Dixon et al., 2012, 2015).

The definition of a TAD and its boundaries depends on the computational algorithms used and the resolution of the HiC experiment and, as such, is somewhat arbitrary. Other types of chromatin domains that represent alternative views of genome compartmentalization have thus been described. SubTADS, for example, are regions of preferential interactions located within a TAD and denote a substructure of these domains (Phillips-Cremins et al., 2013). On a larger scale, TADs can interact with each other, organizing themselves into metaTADs if significant inter-TAD interactions occur in a given cell type or tissue (Fraser et al., 2015). Interestingly, subTADs and metaTADs display a more tissue-specific behavior than TADs (Dixon et al., 2016; Zhan et al., 2017). Loop domains, which correspond to genomic regions of preferential interactions with interaction peaks at their corners (see Glossary, Box 1), have also been described (Rao et al., 2014). Such domains and interactions occur between the boundaries of a TAD, but also within TADs, most of the time involving binding sites for the transcriptional regulator CTCF (see Glossary, Box 1) in convergent orientation. Insulated neighborhoods (see Glossary, Box 1) are specific types of loop domains that are defined by chromatin loops sustained by a CTCF-CTCF homodimer co-bound by cohesin (see Glossary, Box 1), and containing at least one gene (Dowen et al., 2014; Hnisz et al., 2016). In comparison with TADs, subTADs, metaTADs, loop domains and insulated neighborhoods are more labile across cell types and change with gene activity and regulation (Dixon et al., 2016). Accordingly, the algorithms that detect these chromatin domains are based on different concepts and thus do not compartmentalize the genome in the same way as TADs (Dixon et al., 2016). Finally, HiC interaction maps have also revealed the existence of another type of higher-order chromatin structure called A/B compartments (see Glossary, Box 1). These compartments correspond to long-range interacting territories consisting of transcriptionally active (A-compartments) and inactive (B-compartments) domains (Lieberman-Aiden et al., 2009). Recently, the 3D reconstruction of chromatin in the nuclei of

DEVELOPMENT
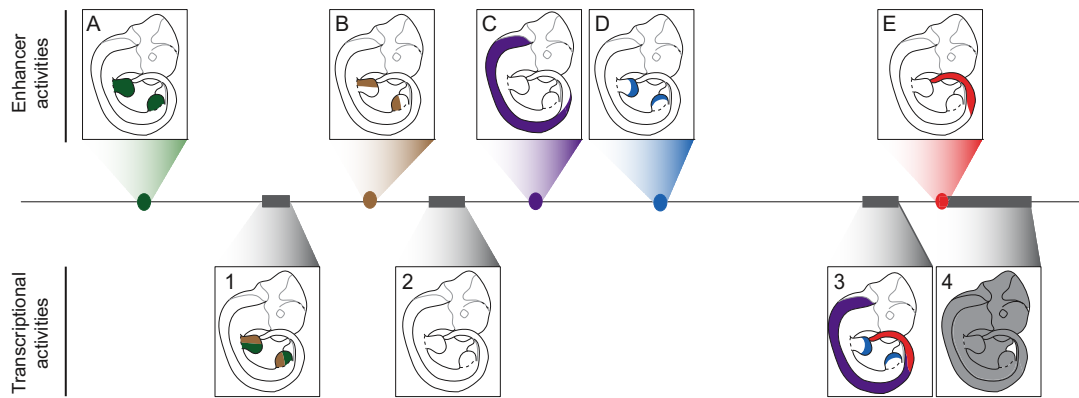
Fig. 1. Enhancers, promoters and their genes together form regulatory landscapes. A hypothetical locus containing five enhancers (colored ovals; A-E) and four genes (gray boxes; 1-4) is shown. The regulatory activities of enhancers A-E in embryonic day 10.5 mouse embryos are represented by colored shading in the diagrams above, and the transcriptional activities of genes 1 to 4 are shown in the diagrams below. A comparison of enhancer activities and gene expression, marked with the same color code, shows that enhancers A and B contribute to gene 1 transcription, and that enhancers C-E contribute to gene 3 transcription. Gene 2 is repressed in all tissues at this embryological stage, whereas gene 4 displays a ubiquitous expression pattern.

single haploid embryonic stem cells (ESCs; see Glossary, Box 1) has shown that A-compartments tend to be found in a ring shape, surrounded by inactive B-compartments, which locate to the periphery and close to the nucleolus (Stevens et al., 2017).

In vertebrates, growing evidence supports the notion that TADs are shaped through the formation of stable DNA loops, called anchor loops, which connect boundary elements together (Rao et al., 2014). Accordingly, TADs correspond to a stable subset of loop domains (Fig. 2). Anchor loops were found to be strongly associated with, and dependent on, the binding of CTCF and the cohesin complex (Sanborn et al., 2015). According to the loop extrusion model, a recently proposed biophysical mechanism for TAD formation, the loading of a cohesin ring around the chromatin progressively extrudes the chromatin until the cohesin ring reaches a 'roadblock', which could be CTCF, to form the observed anchor loops (Sanborn et al., 2015; Fudenberg et al., 2016). According to this model, the constant extruding process is responsible for the frequent mutual interaction of all the sequences located within the TAD. The importance of CTCF and cohesin has also recently been underlined by the finding that the depletion of one of these components leads to a progressive loss of TAD structure, although other chromatin features, such as A/B compartments, are retained (Schwarzer et al., 2016 preprint; Nora et al., 2017). Moreover, two studies have shown that transcription can relocate cohesin over long distances and that its chromatin loading/unloading rate controls the extension of chromatin loops (Busslinger et al., 2017; Haarhuis et al., 2017).

TADs are believed to function as a scaffold for enhancer-promoter interactions (Fig. 2), guiding them in space and allowing them to interact with each other (de Laat and Duboule, 2013). Accordingly, it was recently proposed that TADs act as a buffer for genomic distances and allow for frequent physical interactions to occur between all of their constituent genomic elements (Symmons et al., 2016). In this latter study, the authors showed that the relative genetic distance between the *Shh* gene and its limb enhancer, the ZRS, has no influence on its transcriptional activity as long as both elements locate in the same TAD. In contrast, when the TAD boundary is relocated to lie in between *Shh* and the ZRS, *Shh* transcription is abolished when the genomic distance from the ZRS is large and is reduced when it is short. This result indicated that boundary elements can indeed act as chromatin insulators by reducing the frequency of chromatin interactions but are not impermeable to contact with genomic regions located outside

of the TAD. This concept is supported by studies of the *Sox9* locus, which showed that the duplication of boundary-containing regions results in the formation of a new TAD that is insulated from its neighbors by the duplicated boundary (Franke et al., 2016).

By combining these structural data with enhancer detection methods (see Box 2) and gene expression datasets, an accurate description of the regulatory architecture underlying particular genes can be obtained. This type of approach was recently instrumental in physically linking hundreds of genes involved in human brain development with regulatory regions under positive selective pressure as well as with non-coding variants involved in schizophrenia (Won et al., 2016). Accordingly, with the understanding of chromatin domains such as TADs, subTADs, loop domains or insulated neighborhoods in a given tissue, it is possible to refine the genomic interval important for the regulation of genes as well as the position of putative enhancer regions. It should also be noted that, aside from chromatin domains, subMb-scale chromatin interactions have been shown to structure the genome in a cell type-specific fashion (Javierre et al., 2016; Phillips-Cremins et al., 2013) and, as such, turn out to play a crucial role in gene regulation.

Functional chromatin interactions are dynamic

Despite TADs being rather invariant chromatin structures, extensive differences have been observed within them among different cell types (Dixon et al., 2015). Alterations in intra-TAD chromatin microarchitecture are best shown using viewpoint-specific, proximity ligation technologies, such as 3C, 4C and capture-C/-HiC, because all regions interacting with a specific defined region can be examined at a higher resolution than with HiC (Simonis et al., 2006; Noordermeer et al., 2011; Hughes et al., 2014; Schoenfelder et al., 2015) (Fig. 3, Box 2). Using these methods, extensive changes in the interactions between enhancers and promoters can be observed among different cell types or during embryonic development (Javierre et al., 2016; Andrey et al., 2017; Freire-Pritchett et al., 2017). The first study that used 3C to assess the chromatin architecture of the mouse β-globin locus showed that long-range interactions occur between the β-globin promoter and a cluster of enhancers called the locus control region (LCR); these interactions occur specifically in the β-globin-expressing erythroid cells, but not in brain cells that do not express it (Tolhuis et al.,

Box 3. Proximity ligation technologies

Proximity ligation technologies are used to estimate the frequency of DNA interactions. They are based on the ability to crosslink two genomic loci that are in close physical proximity to each other in the nuclear space (Sati and Cavalli, 2017). The crosslinked chromatin is then subjected to restriction enzyme digestion, which fragments the genomic DNA. The digested chromatin is re-ligated so that regions that are crosslinked together (intra-molecular re-ligation) are attached to each other. This process produces libraries of chimeric DNA products, including adjacent restriction fragments, which are normally found at different positions along the linear genome. Several variations of this technologies exist and include:

Chromosome conformation capture (3C). This approach allows any candidate chimeric product generated from the crosslinking step described above to be quantified using quantitative PCR and specific primers amplifying the product junction (Simonis et al., 2006).

Circular chromosome conformation capture (4C). In this approach, libraries of chimeric DNA products are digested with a second restriction enzyme and further re-ligated together to decrease their molecular size. Using circular PCR, all chimeric products ligated with a desired viewpoint (i.e. a specific restriction fragment) are amplified and sequenced, generating an interaction profile measuring all DNA fragments interacting with the viewpoint (Noordermeer et al., 2011).

Carbon copy chromosome conformation capture (5C). In this approach, primers are designed over a defined genomic interval to amplify many possible ligation products. The amplified library is then sequenced or hybridized to a microarray in order to measure the interaction frequencies of all the tested fragments. Ultimately, an interaction map spanning the defined genomic interval is produced that displays a 'many versus many' interaction map (Dostie and Dekker, 2007).

HiC. In the HiC technique, after the initial restriction enzyme step, the ends are filled in with a biotin-marked nucleotide and subsequently re-ligated. A streptavidin pull-down step is used to enrich for the chimeric products, which are then sequenced. Sequencing of the library produces a genome-wide interaction map of 'all versus all' restriction fragments (Lieberman-Aiden et al., 2009; Rao et al., 2014).

Capture-C (no biotin fill in) and capture-HiC (with biotin fill in). In capture-C, the library of chimeric DNA products is sheared and hybridized to RNA baits, which consist of 100- to 150-bp-long RNA oligonucleotides attached to beads, to enrich for specific loci; the hybridized products are then sequenced in a paired-end fashion to determine the chimeric products (Hughes et al., 2014). The RNA-enriched regions can either be specific genomic viewpoints in order to produce a 'one versus all' 4C-like interaction track or entire loci to produce targeted 'many versus many' interaction maps (Franke et al., 2016; Andrey et al., 2017). In capture-HiC, specific loci or viewpoints are enriched from a HiC library, also using RNA baits (Schoenfelder et al., 2015). These methodologies allow one to either parallelize the production of thousands of interaction tracks from specific viewpoints or produce targeted HiC maps with a much lower sequencing effort compared with HiC.

2002). At other mouse genomic loci, such as at the *HoxD* cluster or at the *Satb1* gene, extensive changes in enhancer-promoter interactions are also observed during development, allowing tissue-specific enhancer-promoter contacts to form (van de Werken et al., 2012; Andrey et al., 2013). Through the use of proximity ligation technologies, the dynamics of chromatin interactions from many specific genomic viewpoints can be simultaneously studied in different tissues or cell types (Simonis et al., 2006; Noordermeer et al., 2011; Hughes et al., 2014; Schoenfelder et al., 2015). For example, by probing the promoter interactome of 17 human primary hematopoietic cell types, one study has shown that most of the scored interactions are cell type or lineage specific and are thus highly dynamic (Javierre et al., 2016). Moreover, by studying hundreds of viewpoints during

mouse limb and midbrain development, two types of chromatin interactions were identified. The first type of contact remained invariable throughout development and between tissue types. The regions involved in these interactions associate with CTCF and cohesin and are likely to play a structural role in organizing the local folding of chromatin in a tissue-independent manner. The second type of chromatin interaction was tissue and/or time point specific. The sites involved in this type of interaction show enrichment for repressive or active functional chromatin marks, indicating that they belong to regulatory regions (Andrey et al., 2017) (Fig. 3). The association between chromatin modifications and facultative chromatin interactions suggests that mechanisms involving the deposition of chromatin marks might mediate this association. Indeed, and as we describe in the section below, Polycomb-repressed regions are spatially clustered together via PRC1 and PRC2 complexes (Denholtz et al., 2013; Joshi et al., 2015; Schoenfelder et al., 2015). Conversely, at transcriptionally active regions, the mediator complex helps to recruit factors, such as the cohesin complex, that bind at promoters and enhancers and enables chromatin interactions between them (Ebmeier and Taatjes, 2010; Kagey et al., 2010; Larivière et al., 2012; Carlsten et al., 2013; Phillips-Cremins et al., 2013).

An important remaining question is whether the observed enhancer-promoter contacts are a cause or consequence of gene regulation and if looping in general is essential and sufficient to activate a gene. Most interestingly, experimentally induced 'forced' looping between enhancers and promoters, outside of their native cellular context, has provided evidence that looping actively controls the accuracy of the onset of gene expression (Deng et al., 2012, 2014). Specifically, the proximity with which enhancers locate to promoters increases the number of alleles that are transcribed per cell (the so-called transcriptional burst fraction), but not the number of RNA molecules produced per transcriptional burst (the burst size) (Bartman et al., 2016). However, in a naturally occurring transcriptional situation, both the burst size and burst fraction are increased, suggesting that the burst size is independent of the spatial proximity between enhancers and promoters. Accordingly, cell fate changes and lineage commitment are associated with large changes in 3D chromatin architecture, and chromatin looping appears to be a driving force in this process. To understand the ground state of this architecture, many studies have thus focused on genome structure in pluripotent cells.

3D genome architecture in pluripotent cells

Pluripotent cells (PCs) share certain features of genomic organization with differentiated cells; they maintain TAD structures as well as long-range A/B compartments that associate with regions of transcriptional activity. However, in contrast to other cell types, PCs display a less compact chromatin organization (Meshorer et al., 2006; Melcer and Meshorer, 2010; Gaspar-Maia et al., 2011). Moreover, long-range contacts between repressed regions are more non-specific in PCs than in differentiated cell types, suggesting that transcriptionally inactive genomic compartments are less strongly established in PCs than in differentiated cells (de Wit et al., 2013).

In PCs, pluripotency gene loci have a tendency to come into contact with each other in *cis* and in *trans* in the nuclear space, somehow sharing a common, cell type-specific, transcriptional machinery (de Wit et al., 2013) (Fig. 4A). Regions bound by pluripotency factors [OCT4 (POU5F1), KLF4, SOX2 and NANOG] were found to colocalize in the nucleus and to form long-range chromatin contacts with each other, indicating that
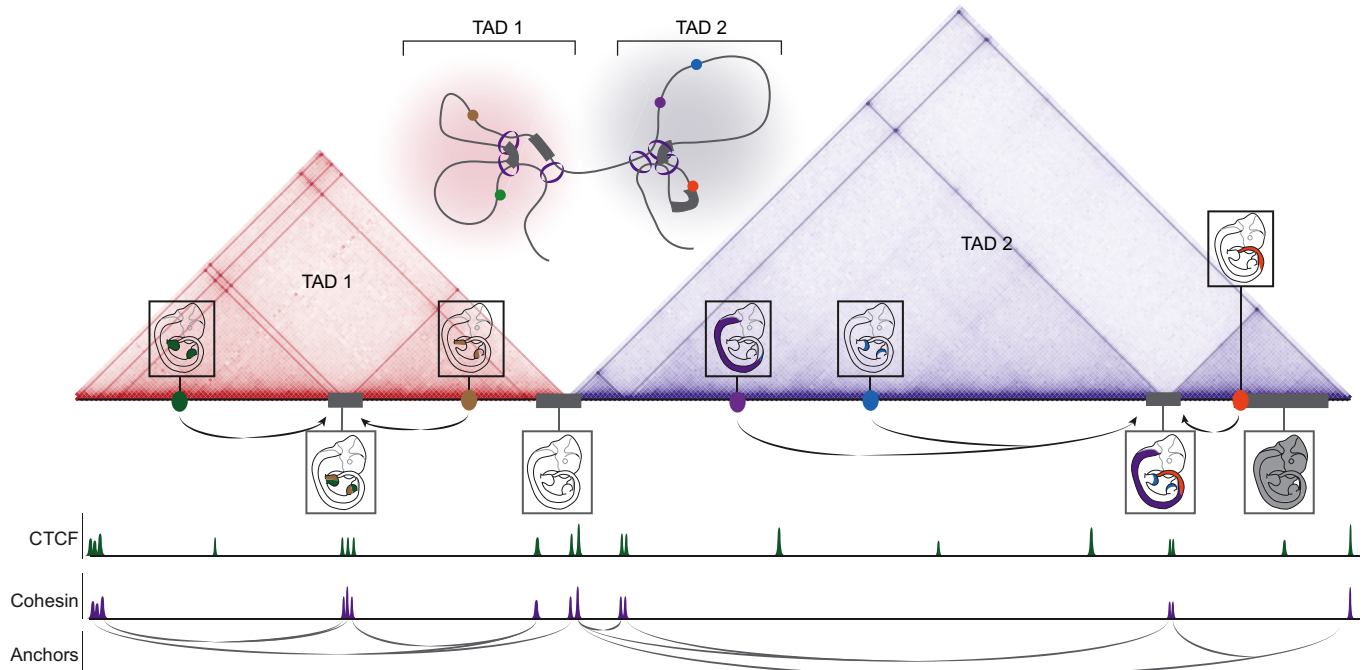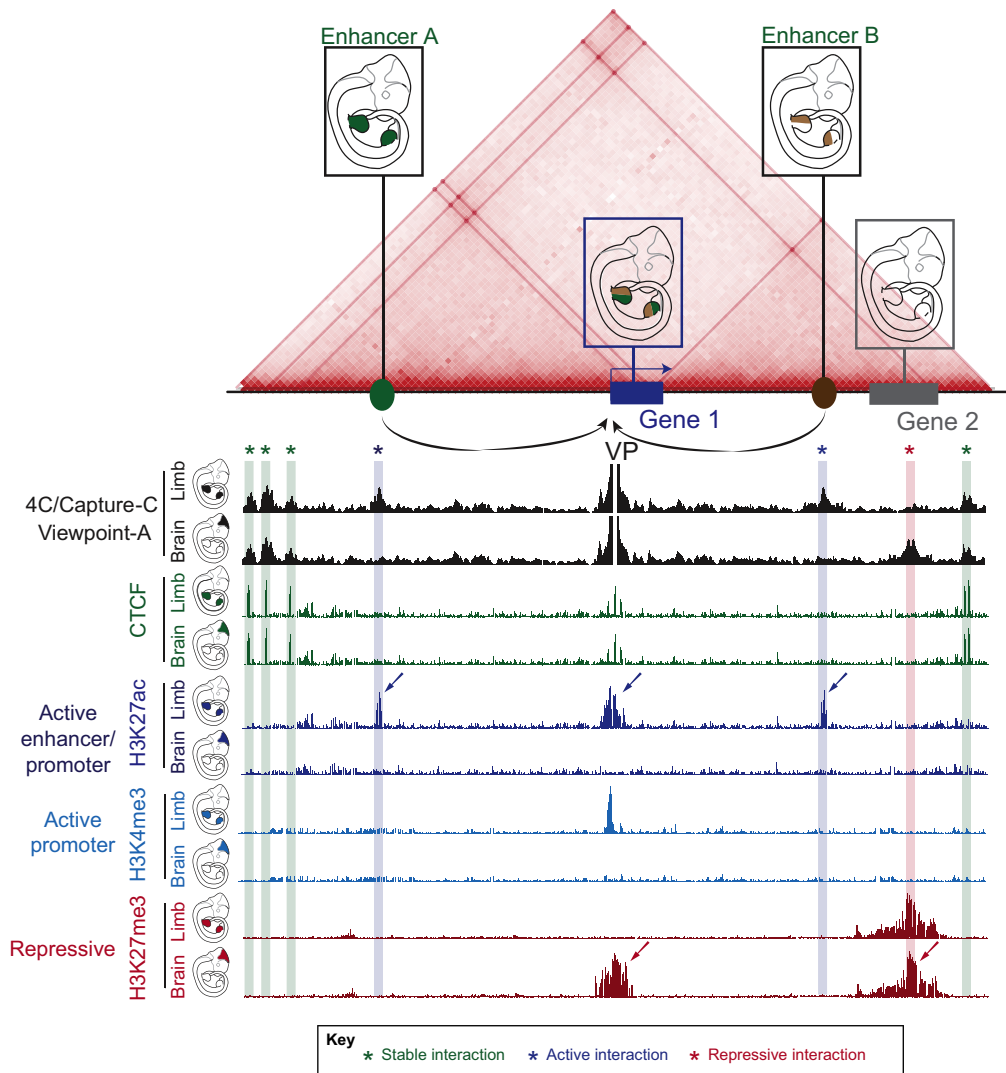
Fig. 2. Topologically associating domains (TADs) form insulated regulatory units. A hypothetical locus containing five enhancers (colored ovals) and four genes (gray boxes) that are organized into two TADs (TAD 1 and TAD 2, depicted by red- and purple-shaded triangles) is shown. A representative chromatin configuration at this locus is shown in between the two TADs, with a cohesin ring surrounding the anchor points. The TADs are sustained by anchor elements (shown as colored lines within the triangles). These elements interact with each other to form loops (indicated by red/purple dots) between distal chromatin regions. CTCF and cohesin (shown as purple rings in the schematic) bind to these elements to form 'loop anchors'. TADs constrain the activities of enhancers and restrict their interactions (shown as black arrows between enhancers) to a limited number of genes. The resulting expression pattern driven by each enhancer in E10.5 mouse embryos is illustrated in boxes. The binding profiles of CTCF (green) and cohesin (purple) along the locus are also shown. Loop anchors are able to form between regions bound by both CTCF and cohesion; accordingly, CTCF sites that are not bound by cohesin do not form loop anchors.

pluripotency factors themselves partially control the genomic architecture of PCs (Denholtz et al., 2013; Bouwman and de Laat, 2015). A study using high-resolution imaging of SOX2-bound enhancers in mouse PCs has revealed their spatial clustering in the nucleus (Liu et al., 2014). In another experimental setting, Wei and colleagues showed that several regions bound by KLF4 are in close spatial proximity to each other in PCs and released upon differentiation or depletion of *Klf4* (Wei et al., 2013). These results were further confirmed by single-cell HiC experiments in haploid mouse ESCs (Stevens et al., 2017). At the intra-TAD level, the formation of chromatin loops was also found to depend on pluripotency factors, which are lost upon differentiation (Kagey et al., 2010; Phillips-Cremins et al., 2013). In particular, several instructive enhancer-promoter interactions, within PC-specific loops, have also been found at the *Sox2* locus (Li et al., 2014). Regions enriched for the binding of Polycomb proteins (see Glossary, Box 1), which are involved in the repression of developmental genes, also frequently interact with each other in PCs, highlighting a role for this protein family in genome organization during pluripotency. Accordingly, when components of the Polycomb PRC1 or PRC2 complexes (see Glossary, Box 1) are inactivated in mouse ESCs, chromatin contacts between Polycomb-repressed genes become altered (Denholtz et al., 2013; Joshi et al., 2015; Schoenfelder et al., 2015).

The establishment of PC-specific chromatin architecture has also been used as a model to study the dynamics of genome organization and the factors that influence this. At various loci, the cohesin and mediator complexes were shown to co-bind at chromatin interaction sites that were also occupied by pluripotency factors (Phillips-Cremins et al., 2013). Upon differentiation, the pluripotency factors are lost and interactions between transcription factor-bound regions are significantly decreased, suggesting that pluripotency factors tether architectural complexes such as the cohesin and mediator complexes (Levasseur et al., 2008; de Wit et al., 2013). In the reverse situation, upon the reprogramming of various cells into induced pluripotent stem cells (iPSCs; see Glossary, Box 1), the PC-specific genome topology is re-established to a large extent, and A/B domains specifically as well as enhancer-promoter interactions at pluripotency loci are very similar to those in ESCs (Beagan et al., 2016; Krijger et al., 2016). It is noteworthy that the maintenance of subMb-scale topological memory of founder cell hallmarks was observed in both studies.

In PCs, as in other cell types, two possible regulatory architectures underlie gene activity: instructive or permissive. An instructive architecture refers to tissue-specific regulatory interactions whereas a permissive one refers to tissue-independent regulatory interactions that are set independently of gene activation (de Laat and Duboule, 2013). An example of permissive architecture in PCs has recently been described at poised enhancers of early differentiation genes. These regions are enriched for H3K27me3 and interact with their putative target genes, prior to their activation, in a PRC2-dependent manner (Cruz-Molina et al., 2017). This suggests that Polycomb interactions might contribute to the formation of a permissive regulatory environment, in which the regulatory contacts of these genes are poised for activation (Fig. 4B). Similarly, several stem cell enhancers are found in tissue-invariant chromatin loops and are

Fig. 3. Dynamic chromatin structure within TADs. The structure of chromatin is dynamic within topologically associating domains (TADs). Active enhancers associate with transcriptionally active promoter regions, which are marked by H3K27ac or H3K4me3 chromatin modifications. By contrast, transcriptionally repressed regions, which are marked by H3K27me3, interact together. The schematic shows a TAD (red-shaded triangle) that encompasses two genes (Gene 1 and Gene 2; shaded blue and gray, respectively). The expression patterns driven by enhancers A and B (green and brown ovals) of Gene 1 in E10.5 mouse embryos are shown inside boxes. Gene 2 is not expressed. The tracks shown beneath the schematic depict the various interactions, binding patterns and modifications seen across this stretch of chromatin. The first tracks show intra-TAD interaction profiles (generated by 4C or capture-C) from the Gene 1 viewpoint (VP), highlighting tissue-independent chromatin contacts with the TAD boundary in both limb and brain tissues, as marked by green vertical lines. These tissue-independent chromatin interactions also associate with the binding of CTCF, as shown in the second set of schematic tracks. In limb tissues, Gene 1 is active, marked by H3K27ac and H3K4me3 (blue arrows, third and fourth sets of schematic tracks, respectively) and its promoter interacts with both active enhancers (A and B), which are also marked by H3K27ac (blue arrows, third schematic track; see blue-shaded, vertical lines). In brain tissue, Gene 1 chromatin is modified by the Polycomb-associated mark H3K27me3 (red arrow, last two schematic tracks) and interacts with Gene 2, which is also modified by H3K27me3 (red arrow, see the inactive interaction red vertical line). Colored asterisks mark stable (green), active (blue) and repressive (red) chromatin interactions within this TAD.

demarcated by the architectural proteins CTCF and cohesin. When these CTCF-binding sites are genetically disrupted using CRISPR/ Cas9 (see Box 4), the misregulation of neighboring genes is observed, suggesting that this permissive environment insulates neighboring loci (Dowen et al., 2014). In summary, whether it is in an instructive or permissive state, the chromatin structure of PCs appears to be flexible and poised for transcriptional activation of early or late differentiation genes.

3D genome architecture during lineage commitment and development

The function and extent of dynamic chromatin interactions are particularly visible during cell differentiation and more generally

during developmental processes. During lineage commitment, extensive switching of chromatin contacts between A and B compartments, which associate with active and inactive transcriptional domains, respectively, is observed. Specifically, around 36% of compartments switch from one type to another as human ESCs differentiate into various lineages (Dixon et al., 2015). Moreover, many pluripotency loci become re-positioned toward the nuclear lamina as they are shut off (Peric-Hupkes et al., 2010). Thus, higher-order genome structure undergoes a widespread remodeling during cell differentiation, which underlies the regulatory decisions associated with cell fate.

As in PCs, the 3D regulatory chromatin architecture that underlies developmental transcriptional regulation is either permissive or
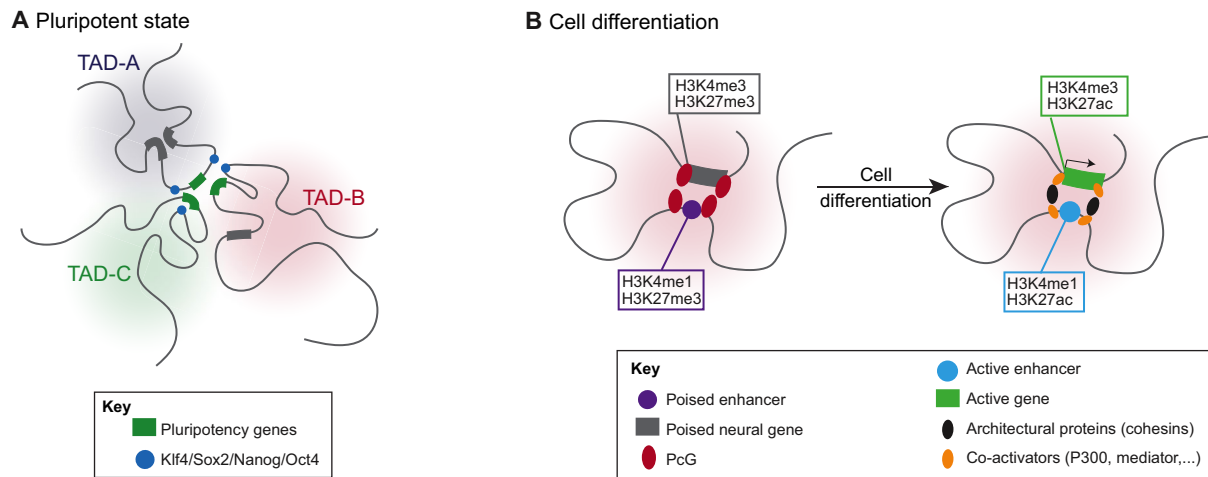
Fig. 4. Organization of the pluripotent genome. (A) Schematic showing how the promoters of pluripotent genes (green), as well as their cis-regulatory elements bound by pluripotency factors (such as Klf4, Sox2, Nanog and Oct4; blue circles) interact with each other in mouse pluripotent stem cells. Three topologically associating domains (TADs A,B and C) promote intra-TAD interactions between regulatory regions and their target promoters by confining local chromatin contacts. (B) The enhancers of genes that require immediate activation in early development, such as anterior neural genes, form loops that connect them to their target promoters through the mediation of Polycomb group (PcG) proteins. This set of genes is poised for transcriptional activation and bears both H3K4me3 and H3K27me3 chromatin marks. Upon differentiation, PcG and H3K27me3 chromatin marks are lost, and H3K27ac is gained both at the promoter and at the enhancer. Therefore, the PcG mediated interaction is lost and replaced by alternative looping mechanisms involving transcriptional co-activators and architectural proteins.

instructive. Accordingly, in a large-scale study of the promoter interactomes of PCs and neuroectodermal cells, around 50% of promoter-interacting regions were re-wired and 50% were retained between both cell types, arguing for extensive chromatin dynamics upon cell differentiation (Freire-Pritchett et al., 2017). Early on in differentiation, transcriptionally active pluripotent genes, such as *Sox2*, *Oct4* and *Nanog*, are repressed and lose their instructive active chromatin interactions with pluripotent enhancers (Phillips-Cremins et al., 2013; Li et al., 2014). In contrast, early differentiation genes, such as early posterior neural genes, which display poised, permissive, PRC2-dependent interactions with their enhancers in PCs, maintain these pre-formed interactions and are therefore rapidly activated upon differentiation (Fig. 3) (Cruz-Molina et al., 2017).

Studies of various specific loci have also provided insights into the dynamic nature of the genome during development. RNA and DNA fluorescence *in situ* hybridization experiments in cells of the mouse posterior limb show that the *Shh* enhancer is in close physical proximity to the *Shh* promoter in those cells that express *Shh* (Amano et al., 2009). Conversely, in anterior limb cells, in which the ZRS enhancer is not active, the *Shh* promoter is nevertheless in close physical proximity, but no *Shh* transcription is observed (Amano et al., 2009; Williamson et al., 2016). Here, a permissive type of 3D structure is at play, and thus the activity of the enhancer alone dictates the expression of the *Shh* gene. Similarly, the transcriptional response of target genes to glucocorticoid receptor (GR) activation has been shown to occur without significant remodeling of pre-formed chromatin contacts, allowing for the rapid onset of gene transcription (Hakim et al., 2011). Interestingly, the mouse *Hoxd13* promoter forms several tissue-non-specific chromatin contacts with distal enhancer regions, as well as a few tissue-specific interactions within the same genomic landscape. Specifically, *Hoxd13* establishes a digit-specific interaction with the digit enhancer *island-3*, and a genital tubercle-specific interaction with the genital tubercle enhancer GT-2 (Montavon et al., 2011; Lonfat et al., 2014). In these cases, as most of the chromatin structure is permissive, some regulatory interactions are formed in a

tissue-specific manner. Other Hoxd genes display a different regulatory architecture. The central mouse Hoxd genes, from *Hoxd9* to *Hoxd11*, regulate limb patterning and shift their contacts between two adjacent TADs. In early and proximal limb tissues, they establish interactions with early enhancers in a telomeric TAD; during later development and in more distal limb regions, these Hoxd genes shift their interactions toward a centromeric TAD that contains digit enhancers. This complex regulatory transition allows for the formation of an intermediary zone, in which low Hoxd expression helps to pattern the wrist (Andrey et al., 2013). HOXA13 and HOXD13 then establish the chromatin structure required for digit-specific patterning (Beccari et al., 2016). At the β-globin and *Satb1* loci, extensive changes in enhancer-promoter interactions occur during erythroid maturation, suggesting that these genes rely on an instructive type of transcriptional onset (Palstra et al., 2003; Vernimmen et al., 2007; van de Werken et al., 2012; Deng et al., 2014). Tissue-specific transcription factors have also been shown to mediate these newly established contacts during lineage commitment. For example, at the β-globin locus, erythroid-specific transcription factors such as EKLF (KLF1), GATA1 and FOG1 (ZFPM1) mediate the cell-specific interactions that occur between the gene locus and its associated enhancers (Drissen et al., 2004; Vakoc et al., 2005).

It is possible that dynamic chromatin interactions act as an active mechanism to control gene transcription rather than representing a passive byproduct of it. Accordingly, the chemical inhibition of transcription at the β-globin locus does not prevent chromatin interactions to occur in a tissue-specific manner with the LCRs, thereby disconnecting transcription from chromatin structure (Palstra et al., 2008). Another study has also shown that the formation of ectopic loops between the LCR and the β-globin gene in the pro-erythroblast cell line GE1, in which β-globin is normally not expressed, results in the strong overexpression of this gene (Deng et al., 2012). The diversity of regulatory architectures observed at different loci has evolved in order to produce a spatially, temporally and lineage-specific expression pattern for the involved genes. Accordingly, and as we discuss below, changes in 3D

chromatin folding, in particular in the structure of TADs, induced by targeted or normal mutations have been shown to affect gene expression and to result in disease or malformation.

Altered TAD structure perturbs enhancer-promoter communication and gene expression: insights into disease

As we have highlighted above, TADs help to shape the overall architecture of the genome and thereby insulate and delineate the extent of regulatory cues. Changes that affect the insulation properties of TADs can thus allow regulatory elements to ectopically contact promoters in neighboring TADs and thereby induce potentially pathogenic gene misexpression. The disruption of TAD structures by deletions, duplications or inversions, collectively called structural variants (SVs), can lead to such effects through enhancer-adoption mechanisms (Fig. 5A-C). Using CRISPR-derived methodologies (see Box 4) to induce SVs, such mutations can be re-engineered in mice (Kraft et al., 2015). Using this approach, it has been shown that a deletion that breaks a TAD boundary and decreases the genomic distance between a set of *Epha4* limb enhancers and the *Pax3* gene leads to the overexpression of *Pax3* in an *Epha4*-like pattern in the mouse limb (Lupiáñez et al., 2015). In this situation, *Pax3* adopts the *Epha4* enhancers and is then misexpressed, leading to a shortening of phalanges. In line with the insulating function of TAD boundaries, a similar effect was not observed when a slightly smaller deletion, leaving the TAD boundary intact, was induced. In addition, an inversion at the *Wnt6/Epha4* locus re-positions the same *Epha4* enhancers in the vicinity of *Wnt6* and induces its ectopic expression in the developing distal limb bud, leading to a specific type of digit malformation, called F-syndrome. Finally, duplications that re-position *Epha4* enhancers in front of the *Ihh* gene induce the overexpression of this gene in distal developing mouse limb buds, leading to polydactyly.

In cancer patients, oncogenes such as *TAL1* or *LMO2* have also been found to be activated as a result of boundary deletion causing the loss of insulated neighborhoods (Hnisz et al., 2016). Most of these spontaneously occurring deletions span boundary regions

of several kilobases and do not precisely pinpoint the underlying set of factors that are responsible for establishing the boundary. As stated above, an important factor implicated in boundary formation is the transcription factor CTCF. Upon the genetic removal of CTCF-bound regions in mouse ESCs, which correspond to domain boundaries, ectopic interactions between genes and enhancers are observed that lead to gene misexpression (Dowen et al., 2014). This suggests that mutations that affect CTCF binding in human patients would also lead to gene misexpression and potentially to disease. Accordingly, in cancer patients with IDH gliomas, the ectopic methylation of a boundary element has been shown to associate with decreased CTCF binding, and with increased inter-TAD interactions and the misexpression of the *PDGFRA* oncogene (Flavahan et al., 2016). However, it is unclear if this mechanism can be applied to other loci, as CTCF was previously shown to bind independently of methylation levels and to trigger demethylase activity in mouse cells (Stadler et al., 2011).

Another type of mechanism has recently been shown to generate pathogenic effects through genomic duplications. The duplication of regions that encompass enhancers as well as a TAD boundary can result in the formation of new chromatin domains, called neo-TADs, that are isolated from the rest of the genome (Franke et al., 2016; Weischenfeldt et al., 2017). If the duplicated region does not contain genes, its regulatory activities remain restricted to the neo-TAD, and are therefore without effect. However, if a gene is included in the duplication, it will adopt the regulatory activities of the neo-TAD thereby acquiring new expression domains (Fig. 5D). This phenomenon can lead to ectopic gene expression, as shown for the *Kcnj2* gene in mice, which adopts a *Sox9* pattern when included in the neo-TAD that contains duplicated *Sox9* enhancers (Franke et al., 2016). This *de novo* expression of *Kcnj2* results in a congenital malformation that is characterized by nail aplasia and short digits (Franke et al., 2016). In this example, the regulation of neighboring genes remains largely unchanged. This phenomenon has also been described at the human *IGF2* locus, where genomic duplications create a new TAD domain that consists of the *IGF2* gene and a colorectal cancer lineage-specific super enhancer, which results in the misregulation of the oncogenic locus (Weischenfeldt et al., 2017).

It can be easily envisioned that the formation of new regulatory domains that are separated from the rest of the genome by boundaries provide an ideal setting in which to acquire new functions in an evolutionary context. These results suggest that the formation and isolation of a newly formed TAD can result in a phenotypic change in an organism that is then directly subjected to selective pressure, without affecting the parent copy of the gene. In that sense, the shifting of TADs and the recombining of regulatory activity with new target genes provides a toolbox of possibilities for how new gene functions can be acquired (Franke et al., 2016).

Conclusions

A crucial advance in our understanding of gene regulation came with the finding that the genome undergoes three-dimensional folding in the nucleus in a genetically determined process that directly influences gene regulation via the formation of chromatin units called TADs. These units directly influence the availability of enhancers for their target genes. The complex interplay between distal genomic sequences that contain regulatory elements and their target genes and promoters has proven to be essential for regulating genes during development and lineage commitment. In particular, the development of new technologies that measure the frequency of DNA interactions has advanced the field, enabling the analyses
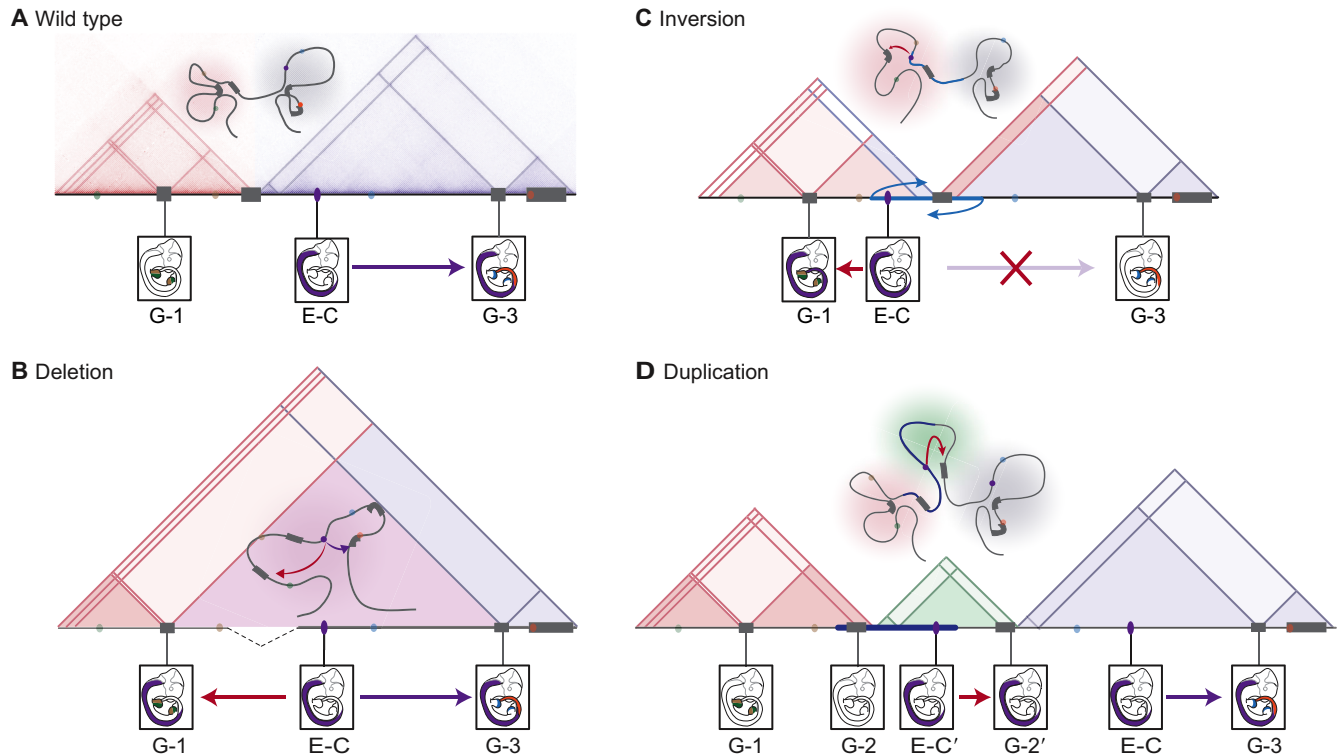
Fig. 5. Structural variants affect TAD structure and enhancer-promoter interactions. Structural variants (SVs) can induce ectopic or loss of gene expression leading to disease. Here, topologically associating domains (TADs) are represented by shaded triangles, genes by gray blocks, and enhancers by colored ovals. Gene expression patterns driven by enhancers in E10.5 mouse embryos are shown. A representative looped chromatin configuration is shown above the TADs. (A) In the wild-type chromatin conformation, Enhancer C (E-C) controls Gene 3 (G-3) in the neural tube. (B) The deletion of a boundary element between the two TADs leads to ectopic contact between Enhancer C (E-C) and Gene 1 (G-1) (red arrow), which results in the ectopic expression of Gene 1 in the neural tube. (C) In the case of an inversion (blue arrows) that leads to the repositioning of functional elements, Enhancer C (E-C) from a neighboring TAD is free to activate Gene 1 (G-1), resulting in the ectopic expression of Gene 1 in the neural tube. The inversion also leaves Enhancer C insulated from its native target Gene 3 (G-3) by a boundary, resulting in the loss of Gene 3 expression in the neural tube. (D) The duplication of a region allows a new chromatin domain (a neo-TAD; green triangle) to form that contains regulatory region(s) and gene(s), which produce new expression patterns. Here, the duplicated Enhancer C′ (E-C′) and Gene 2′ (G-2′), which both locate to the insulated neo-TAD, produce the ectopic expression of Gene 2′ in the neural tube.

of multiple genomic regions at high resolution, even in single cells. Moreover, the recently developed genome architecture mapping (GAM) approach, which does not rely on crosslinking or ligation, has revealed numerous three-way interactions between chromatin regions and will increase our capacity to investigate such complex chromatin interactions (Beagrie et al., 2017). Furthermore, the development of direct visualization technologies that rely on super-resolution microscopy techniques, especially *in vivo*, should provide novel insights into the allele-to-allele variability of dynamic chromatin interactions.

The role of the 3D genome architecture during development or lineage commitment can be described in two ways. In the first rather structural view, the genome is divided into chromatin domains that have been described in different ways, such as TADs, loop domains or insulated neighborhoods. One interesting feature of TADs is their maintenance during the oocyte-to-zygote transition in mice or their establishment prior to genome transcriptional activation in *Drosophila*, which suggests that they are formed prior to the transcriptional function they control (Flyamer et al., 2017; Hug et al., 2017). Chromatin domains, and in particular TADs, are in this respect believed to act by delineating the regions scanned by enhancer and promoters, independently of tissues or cell types. The second view takes into account the extensive chromatin dynamics that occur during cell differentiation

and development, within and in between domains. This is best exemplified by the recently characterized interactome footprints, which are sufficient to identify specific cell types (Javierre et al., 2016). It will be important to understand which factors induce the formation of these variable chromatin structures. Obviously, tissue-specific transcription factors play a key role in producing them but we also need to identify the mechanisms through which they act. Changes in chromatin modifications, which directly derive from the binding of transcription factors, can also be associated with the dynamics of chromatin interactions (Javierre et al., 2016; Andrey et al., 2017). Thus, transcription factors might act by recruiting other factors to the chromatin, which are linked to chromatin modifications and which themselves trigger the looping.

Future challenges will involve the dissection of the biochemical mechanisms underlying the formation of permissive or instructive chromatin loops. These findings, as well as progress in modeling technologies, will help to improve predictions of the architectural outcome and thus pathogenic effect of mutations involved in various conditions, including congenital malformations and cancers. The quantification of chromatin architecture and associated dynamics in heterogeneous tissue will benefit a lot from purer cell populations, obtained by, for example, fluorescence-activated cell sorting technologies or high-resolution microscopy.

DEVELOPMENT

Together, these studies will also help establish the causal link between chromatin interactions and gene expression. Important studies have already paved the way to the modulation of tissue-specific loops, either by activating them prior to their normal activation or by changing their interaction partners in cell culture systems (Deng et al., 2012, 2014; Bartman et al., 2016). However, the regulatory role that these dynamic chromatin interactions play in establishing the tightly regulated expression patterns of development genes *in vivo*, which ultimately control the morphogenesis of organs and structures, remains to be determined.

References

Akhtar, W., de Jong, J., Pindyurin, A. V., Pagie, L., Meuleman, W., de Ridder, J., Berns, A., Wessels, L. F. A., van Lohuizen, M. and van Steensel, B. (2013). Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell* **154**, 914-927.

Amano, T., Sagai, T., Tanabe, H., Mizushina, Y., Nakazawa, H. and Shiroishi, T. (2009). Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* **16**, 47-57.

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T. et al. (2014). An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455-461.

Andrey, G. and Spielmann, M. (2017). CRISPR/Cas9 genome editing in embryonic stem cells. *Methods Mol. Biol.* **1468**, 221-234.

Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F. and Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* **340**, 1234167.

Andrey, G., Schöpflin, R., Jerković, I., Heinrich, V., Ibrahim, D. M., Paliou, C., Hochradel, M., Timmermann, B., Haas, S., Vingron, M. et al. (2017). Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding. *Genome Res.* **27**, 223-233.

Araya, C. L., Kawli, T., Kundaje, A., Jiang, L., Wu, B., Vafeados, D., Terrell, R., Weissdepp, P., Gevirtzman, L., Mace, D. et al. (2014). Regulatory analysis of the C. elegans genome with spatiotemporal resolution. *Nature* **512**, 400-405.

Arnold, C. D., Gerlach, D., Stelzer, C., Boryn, L. M., Rath, M. and Stark, A. (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074-1077.

Arnold, C. D., Zabidi, M. A., Pagani, M., Rath, M., Schernhuber, K., Kazmar, T. and Stark, A. (2017). Genome-wide assessment of sequence-intrinsic enhancer responsiveness at single-base-pair resolution. *Nat. Biotechnol.* **35**, 136-144.

Banerji, J., Rusconi, S. and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299-308.

Bartman, C. R., Hsu, S. C., Hsiung, C. C.-S., Raj, A. and Blobel, G. A. (2016). Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol. Cell* **62**, 237-247.

Beagan, J. A., Gilgenast, T. G., Kim, J., Plona, Z., Norton, H. K., Hu, G., Hsu, S. C., Shields, E. J., Lyu, X., Apostolou, E. et al. (2016). Local genome topology can exhibit an incompletely rewired 3D-folding state during somatic cell reprogramming. *Cell Stem Cell* **18**, 611-624.

Beagrie, R. A., Scialdone, A., Schueler, M., Kraemer, D. C. A., Chotalia, M., Xie, S. Q., Barbieri, M., de Santiago, I., Lavitas, L.-M., Branco, M. R. et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519-524.

Beccari, L., Yakushiji-Kaminatsui, N., Woltering, J. M., Necsulea, A., Lonfat, N., Rodriguez-Carballo, E., Mascrez, B., Yamamoto, S., Kuroiwa, A. and Duboule, D. (2016). A role for HOX13 proteins in the regulatory switch between TADs at the HoxD locus. *Genes Dev.* **30**, 1172-1186.

Birling, M.-C., Schaeffer, L., André, P., Lindner, L., Maréchal, D., Ayadi, A., Sorg, T., Pavlovic, G. and Hérault, Y. (2017). Efficient and rapid generation of large genomic variants in rats and mice using CRISMERE. *Sci. Rep.* **7**, 43331.

Bonev, B. and Cavalli, G. (2016). Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661-678.

Bothma, J. P., Garcia, H. G., Ng, S., Perry, M. W., Gregor, T. and Levine, M. (2015). Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. *eLife* **4**.

Bouwman, B. A. M. and de Laat, W. (2015). Architectural hallmarks of the pluripotent genome. *FEBS Lett.* **589**, 2905-2913.

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. and Greenleaf, W. J. (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.

Busslinger, G. A., Stocsits, R. R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N. and Peters, J.-M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**, 503-507.

Carlsten, J. O. P., Zhu, X. and Gustafsson, C. M. (2013). The multitalented Mediator complex. *Trends Biochem. Sci.* **38**, 531-537.

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Jr Shapiro, M. D., Brady, S. D., Southwick, A. M., Absher, D. M., Grimwood, J., Schmutz, J. et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**, 302-305.

Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A. et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **107**, 21931-21936.

Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R., van Ijcken, W. F., Grosveld, F., Frommolt, P., Bazzi, H. et al. (2017). PRC2 facilitates the regulatory topology required for poised enhancer function during pluripotent stem cell differentiation. *Cell Stem Cell* **20**, 689-705.e9.

de Laat, W. and Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499-506.

de Wit, E., Bouwman, B. A. M., Zhu, Y., Klous, P., Splinter, E., Verstegen, M. J. A. M., Krijger, P. H. L., Festuccia, N., Nora, E. P., Welling, M. et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* **501**, 227-231.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P. D., Dean, A. and Blobel, G. A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233-1244.

Deng, W., Rupon, J. W., Krivega, I., Breda, L., Motta, I., Jahn, K. S., Reik, A., Gregory, P. D., Rivella, S., Dean, A. et al. (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**, 849-860.

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M. and Plath, K. (2013). Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* **13**, 602-616.

Di Croce, L. and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* **20**, 1147-1155.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S. and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380.

Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J. E., Lee, A. Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W. et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331-336.

Dixon, J. R., Gorkin, D. U. and Ren, B. (2016). Chromatin domains: the unit of chromosome organization. *Mol. Cell* **62**, 668-680.

Dostie, J. and Dekker, J. (2007). Mapping networks of physical interactions between genomic elements using 5C technology. *Nat. Protoc.* **2**, 988-1002.

Dowen, J. M., Fan, Z. P., Hnisz, D., Ren, G., Abraham, B. J., Zhang, L. N., Weintraub, A. S., Schuijers, J., Lee, T. I., Zhao, K. et al. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374-387.

Drissen, R., Palstra, R. J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S. and de Laat, W. (2004). The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev.* **18**, 2485-2490.

Ebmeier, C. C. and Taatjes, D. J. (2010). Activator-Mediator binding regulates Mediator-cofactor interactions. *Proc. Natl. Acad. Sci. USA* **107**, 11283-11288.

Flavahan, W. A., Drier, Y., Liau, B. B., Gillespie, S. M., Venteicher, A. S., Stemmer-Rachamimov, A. O., Suvà, M. L. and Bernstein, B. E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110-114.

Flyamer, I. M., Gassler, J., Imakaev, M., Brandão, H. B., Ulianov, S. V., Abdennur, N., Razin, S. V., Mirny, L. A. and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. *Nature* **544**, 110-114.

Franke, M., Ibrahim, D. M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.-L. et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265-269.

Fraser, J., Ferrai, C., Chiariello, A. M., Schueler, M., Rito, T., Laudanno, G., Barbieri, M., Moore, B. L., Kraemer, D. C., Aitken, S. et al. (2015). Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol. Syst. Biol.* **11**, 852.

DEVELOPMENT

Freire-Pritchett, P., Schoenfelder, S., Varnai, C., Wingett, S. W., Cairns, J., Collier, A. J., Garcia-Vilchez, R., Furlan-Magaril, M., Osborne, C. S., Fraser, P. et al. (2017). Global reorganisation of cis-regulatory units upon lineage commitment of human embryonic stem cells. *eLife* **6**, e21926.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N. and Mirny, L. A. (2016). Formation of chromosomal domains by loop extrusion. *Cell Rep* **15**, 2038-2049.

Gaspar-Maia, A., Alajem, A., Meshorer, E. and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. *Nat. Rev. Mol. Cell Biol.* **12**, 36-47.

Haarhuis, J. H. I., van der Weide, R. H., Blomen, V. A., Yanez-Cuna, J. O., Amendola, M., van Ruiten, M. S., Krijger, P. H. L., Teunissen, H., Medema, R. H., van Steensel, B. et al. (2017). The Cohesin release factor WAPL restricts chromatin loop extension. *Cell* **169**, 693-707.e614.

Hakim, O., Sung, M.-H., Voss, T. C., Splinter, E., John, S., Sabo, P. J., Thurman, R. E., Stamatoyannopoulos, J. A., de Laat, W. and Hager, G. L. (2011). Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. *Genome Res.* **21**, 697-706.

Hay, D., Hughes, J. R., Babbs, C., Davies, J. O. J., Graham, B. J., Hanssen, L. L. P., Kassouf, M. T., Oudelaar, A. M., Sharpe, J. A., Suciu, M. C. et al. (2016). Genetic dissection of the alpha-globin super-enhancer in vivo. *Nat. Genet.* **48**, 895-903.

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W. et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-112.

Hnisz, D., Day, D. S. and Young, R. A. (2016). Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* **167**, 1188-1200.

Hug, C. B., Grimaldi, A. G., Kruse, K. and Vaquerizas, J. M. (2017). Chromatin architecture emerges during zygotic genome activation independent of transcription. *Cell* **169**, 216-228.e219.

Hughes, J. R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R. and Higgs, D. R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* **46**, 205-212.

Javierre, B. M., Burren, O. S., Wilder, S. P., Kreuzhuber, R., Hill, S. M., Sewitz, S., Cairns, J., Wingett, S. W., Varnai, C., Thiecke, M. J. et al. (2016). Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* **167**, 1369-1384.e1319.

Joshi, O., Wang, S.-Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P. J., Habibi, E., Shaik, J., Saeed, S., Handoko, L. et al. (2015). Dynamic reorganization of extremely long-range promoter-promoter interactions between two states of pluripotency. *Cell Stem Cell* **17**, 748-757.

Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S. et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-435.

Kondo, T. and Duboule, D. (1999). Breaking colinearity in the mouse HoxD complex. *Cell* **97**, 407-417.

Kraft, K., Geuer, S., Will, A. J., Chan, W. L., Paliou, C., Borschiwer, M., Harabula, I., Wittler, L., Franke, M., Ibrahim, D. M. et al. (2015). Deletions, inversions, duplications: engineering of structural variants using CRISPR/Cas in mice. *Cell Rep.* **10**, 833-839.

Krijger, P. H. L., Di Stefano, B., de Wit, E., Limone, F., van Oevelen, C., de Laat, W. and Graf, T. (2016). Cell-of-origin-specific 3D genome structure acquired during somatic cell reprogramming. *Cell Stem Cell* **18**, 597-610.

Larivière, L., Plaschka, C., Seizl, M., Wenzeck, L., Kurth, F. and Cramer, P. (2012). Structure of the Mediator head module. *Nature* **492**, 448-451.

Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E. and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725-1735.

Levasseur, D. N., Wang, J., Dorschner, M. O., Stamatoyannopoulos, J. A. and Orkin, S. H. (2008). Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. *Genes Dev.* **22**, 575-580.

Li, Y., Rivera, C. M., Ishii, H., Jin, F., Selvaraj, S., Lee, A. Y., Dixon, J. R. and Ren, B. (2014). CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells. *PLoS ONE* **9**, e114485.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O. et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293.

Liu, Z., Legant, W. R., Chen, B. C., Li, L., Grimm, J. B., Lavis, L. D., Betzig, E. and Tjian, R. (2014). 3D imaging of Sox2 enhancer clusters in embryonic stem cells. *eLife* **3**, e04236.

Lonfat, N., Montavon, T., Darbellay, F., Gitto, S. and Duboule, D. (2014). Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. *Science* **346**, 1004-1006.

Long, H. K., Prescott, S. L. and Wysocka, J. (2016). Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170-1187.

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M., Laxova, R. et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012-1025.

Lupiáñez, D. G., Spielmann, M. and Mundlos, S. (2016). Breaking TADs: how alterations of chromatin domains result in disease. *Trends Genet.* **32**, 225-237.

Manning, L., Heckscher, E. S., Purice, M. D., Roberts, J., Bennett, A. L., Kroll, J. R., Pollard, J. L., Strader, M. E., Lupton, J. R., Dyukareva, A. V. et al. (2012). A resource for manipulating gene expression and analyzing cis-regulatory modules in the Drosophila CNS. *Cell Rep.* **2**, 1002-1013.

Margueron, R. and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343-349.

Marinić, M., Aktas, T., Ruf, S. and Spitz, F. (2013). An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**, 530-542.

Melcer, S. and Meshorer, E. (2010). Chromatin plasticity in pluripotent cells. *Essays Biochem.* **48**, 245-262.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P. J., Brown, D. T. and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell* **10**, 105-116.

Montavon, T., Le Garrec, J.-F., Kerszberg, M. and Duboule, D. (2008). Modeling Hox gene regulation in digits: reverse collinearity and the molecular origin of thumbness. *Genes Dev.* **22**, 346-359.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F. and Duboule, D. (2011). A regulatory archipelago controls Hox genes transcription in digits. *Cell* **147**, 1132-1145.

Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W. and Duboule, D. (2011). The dynamic architecture of Hox gene clusters. *Science* **334**, 222-225.

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N. L., Meisig, J., Sedat, J. et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-385.

Nora, E. P., Goloborodko, A., Valton, A. L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A. and Bruneau, B. G. (2017). Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930-944.e922.

Palstra, R.-J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F. and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* **35**, 190-194.

Palstra, R.-J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B. and de Laat, W. (2008). Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. *PLoS ONE* **3**, e1661.

Pennacchio, L. A., Ahituv, N., Moses, A. M., Prabhakar, S., Nobrega, M. A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K. D. et al. (2006). In vivo enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499-502.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W. M., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R. M., van Lohuizen, M. et al. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603-613.

Petit, F., Sears, K. E. and Ahituv, N. (2017). Limb development: a paradigm of gene regulation. *Nat. Rev. Genet.* **18**, 245-258.

Phillips-Cremins, J. E., Sauria, M. E. G., Sanyal, A., Gerasimova, T. I., Lajoie, B. R., Bell, J. S. K., Ong, C.-T., Hookway, T. A., Guo, C., Sun, Y. et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* **153**, 1281-1295.

Pope, B. D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D. L., Wang, Y., Hansen, R. S., Canfield, T. K. et al. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402-405.

Pradeepa, M. M., Grimes, G. R., Kumar, Y., Olley, G., Taylor, G. C. A., Schneider, R. and Bickmore, W. A. (2016). Histone H3 globular domain acetylation identifies a new class of enhancers. *Nat. Genet.* **48**, 681-686.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A. and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279-283.

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S. et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680.

Ruf, S., Symmons, O., Uslu, V. V., Dolle, D., Hot, C., Ettwiller, L. and Spitz, F. (2011). Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nat. Genet.* **43**, 379-386.

Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M. and Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* **132**, 797-803.

Sanborn, A. L., Rao, S. S. P., Huang, S.-C., Durand, N. C., Huntley, M. H., Jewett, A. I., Bochkov, I. D., Chinnappan, D., Cutkosky, A., Li, J. et al. (2015).

Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. USA* **112**, E6456-E6465.

Sanyal, A., Lajoie, B. R., Jain, G. and Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature* **489**, 109-113.

Sati, S. and Cavalli, G. (2017). Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* **126**, 33-44.

Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.-M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M. et al. (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* **47**, 1179-1186.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N. A., Huber, W., Haering, C., Mirny, L. et al. (2016). Two independent modes of chromosome organization are revealed by cohesin removal. *bioRxiv*.

Shlyueva, D., Stampfel, G. and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* **15**, 272-286.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**, 1348-1354.

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**, pdb prot5384.

Spitz, F. and Furlong, E. E. M. (2012). Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613-626.

Stadler, M. B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E. J., Gaidatzis, D. et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490-495.

Stevens, T. J., Lando, D., Basu, S., Atkinson, L. P., Cao, Y., Lee, S. F., Leeb, M., Wohlfahrt, K. J., Boucher, W., O'Shaughnessy-Kirwan, A. et al. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59-64.

Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W. and Spitz, F. (2016). The Shh topological domain facilitates the action of remote enhancers by reducing the effects of genomic distances. *Dev. Cell* **39**, 529-543.

Thierion, E., Le Men, J., Collombet, S., Hernandez, C., Coulpier, F., Torbey, P., Thomas-Chollier, M., Noordermeer, D., Charnay, P. and Gilardi-Hebenstreit, P. (2017). Krox20 hindbrain regulation incorporates multiple modes of cooperation between cis-acting elements. *PLoS Genet.* **13**, e1006903.

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F. and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* **10**, 1453-1465.

Tsujimura, T., Klein, F. A., Langenfeld, K., Glaser, J., Huber, W. and Spitz, F. (2015). A discrete transition zone organizes the topological and regulatory autonomy of the adjacent tfap2c and bmp7 genes. *PLoS Genet.* **11**, e1004897.

Vakoc, C. R., Mandat, S. A., Olenchock, B. A. and Blobel, G. A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell* **19**, 381-391.

van de Werken, H. J. G., Landan, G., Holwerda, S. J. B., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Öz, Y., Bouwman, B. A. M. et al. (2012). Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969-972.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J. A., Wood, W. G. and Higgs, D. R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* **26**, 2041-2051.

Visel, A., Minovitsky, S., Dubchak, I. and Pennacchio, L. A. (2007). VISTA Enhancer Browser–a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35**, D88-D92.

Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F. and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918.

Wei, Z., Gao, F., Kim, S., Yang, H., Lyu, J., An, W., Wang, K. and Lu, W. (2013). Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. *Cell Stem Cell* **13**, 36-47.

Weischenfeldt, J., Dubash, T., Drainas, A. P., Mardin, B. R., Chen, Y., Stütz, A. M., Waszak, S. M., Bosco, G., Halvorsen, A. R., Raeder, B. et al. (2017). Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat. Genet.* **49**, 65-74.

Wijchers, P. J., Geeven, G., Eyres, M., Bergsma, A. J., Janssen, M., Verstegen, M., Zhu, Y., Schell, Y., Vermeulen, C., de Wit, E. et al. (2015). Characterization and dynamics of pericentromere-associated domains in mice. *Genome Res.* **25**, 958-969.

Will, A. J., Cova, G., Osterwalder, M., Chan, W.-L., Wittler, L., Brieske, N., Heinrich, V., de Villartay, J.-P., Vingron, M., Klopocki, E. et al. (2017). Composition and dosage of a multipartite enhancer cluster control developmental expression of *Ihh* (*Indian hedgehog*). *Nat. Genet.* (in press).

Williamson, I., Lettice, L. A., Hill, R. E. and Bickmore, W. A. (2016). Shh and ZRS enhancer colocalisation is specific to the zone of polarising activity. *Development* **143**, 2994-3001.

Won, H., de la Torre-Ubieta, L., Stein, J. L., Parikshak, N. N., Huang, J., Opland, C. K., Gandal, M. J., Sutton, G. J., Hormozdiari, F., Lu, D. et al. (2016). Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523-527.

Xie, S., Duan, J., Li, B., Zhou, P. and Hon, G. C. (2017). Multiplexed engineering and analysis of combinatorial enhancer activity in single cells. *Mol. Cell* **66**, 285-299.e285.

Zabidi, M. A., Arnold, C. D., Schernhuber, K., Pagani, M., Rath, M., Frank, O. and Stark, A. (2015). Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556-559.

Zeller, R., López-Ríos, J. and Zuniga, A. (2009). Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nat. Rev. Genet.* **10**, 845-858.

Zhan, Y., Mariani, L., Barozzi, I., Schulz, E. G., Blüthgen, N., Stadler, M., Tiana, G. and Giorgetti, L. (2017). Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome Res.* **27**, 479-490.

DEVELOPMENT