The selection and function of cell type-specific enhancers

Sven Heinz[1], Casey E. Romanoski[2], Christopher Benner[1] and Christopher K. Glass[2,3]

Abstract | The human body contains several hundred cell types, all of which share the same genome. In metazoans, much of the regulatory code that drives cell type-specific gene expression is located in distal elements called enhancers. Although mammalian genomes contain millions of potential enhancers, only a small subset of them is active in a given cell type. Cell type-specific enhancer selection involves the binding of lineage-determining transcription factors that prime enhancers. Signal-dependent transcription factors bind to primed enhancers, which enables these broadly expressed factors to regulate gene expression in a cell type-specific manner. The expression of genes that specify cell type identity and function is associated with densely spaced clusters of active enhancers known as super-enhancers. The functions of enhancers and super-enhancers are influenced by, and affect, higher-order genomic organization.

The molecular mechanisms that enable and mediate cell type-specific transcriptional responses to intracellular and extracellular cues remain poorly understood. Early experiments indicated that sequences far away from gene promoters are often required to regulate cell type-specific transcription[1]. Such genetic elements are termed enhancers and were initially functionally defined as DNA sequences that have the potential to enhance basal transcription levels from gene promoters and transcriptional start sites (TSSs)[1] at distances ranging from hundreds of bases to megabases[2]. Recent genome-wide transcription factor-binding studies indicated that the majority of transcription factor binding sites are found in distal locations that frequently exhibit enhancer function[3–9], which is consistent with the profound role of enhancers in shaping signal-dependent transcriptional responses[10–12].

When cell signalling induces an increase in the nuclear concentration and DNA binding of transcription factors, as occurs following the activation of steroid hormone receptors and nuclear factor-κB (NF-κB), the great majority of binding events typically occur at genomic locations that already exhibit enhancer-like histone modifications and binding of other transcription factors[5,6]. As the complement of active *cis*-regulatory elements is different across cell types, these findings introduced the notion that pre-existing sets of enhancers are primarily responsible for cell type-specific gene expression and responses to external stimuli[13–15]. The annotation of epigenetic features that are associated with enhancers in many different cell lines, primary cells and tissues by the Encyclopedia of DNA Elements (ENCODE) consortium provided evidence for the use of several hundreds of thousands of such elements in the human genome[16], which greatly exceeds the number of genes that encode mRNAs or long intergenic non-coding RNAs (lincRNAs). This raised the question of how the correct subsets of enhancers are selected from the large repertoire of potential enhancers in each particular cell type.

Here, we review recent findings on the selection and function of enhancers that specify cell identity and that underlie the distinctive responses of cells to intracellular and extracellular signals. We discuss the collaborative and hierarchical binding of transcription factors to DNA in the context of chromatin, which orchestrates enhancer selection and priming, and the transformation of chromatin from a silent, primed or poised state to one that actively supports transcription. We conclude with a discussion of the 3D organization of enhancers in the nucleus and its importance for their function.

Enhancer characteristics

Genomic regions that function as transcriptional enhancers are enriched in closely spaced recognition motifs for sequence-specific transcription factors. Enhancer activation begins with the binding of transcription factors and local nucleosome remodelling. Recent genome-wide studies of nucleosome remodelling during differentiation of embryonic stem cells (ESCs)

[1]*Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, San Diego, California 92037, USA*
[2]*Department of Cellular and Molecular Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, San Diego, California 92093, USA.*
[3] *Department of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, San Diego, California 92093, USA.*
Correspondence to C.K.G.
e-mail: ckg@ucsd.edu
doi:10.1038/nrm3949
Published online
4 February 2015

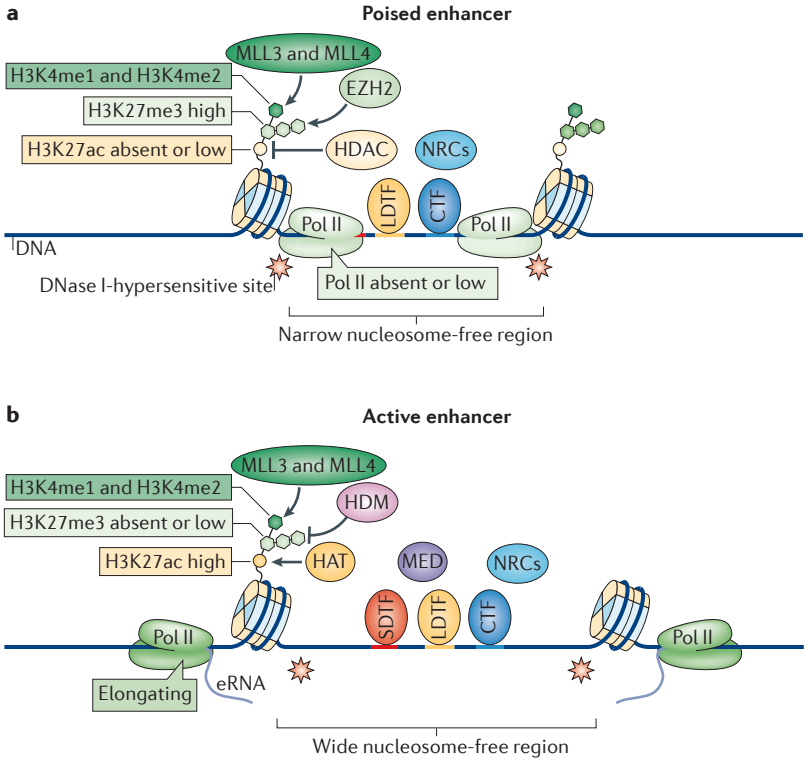

Figure 1 | **The anatomies of poised and active enhancers.** The characteristic features of poised and active enhancers are shown, including the binding of lineage-determining transcription factors (LDTFs) and collaborating transcription factors (CTFs) to closely spaced recognition motifs (yellow and blue sites, respectively) on the DNA. **a** | The binding of these factors in concert with nucleosome-remodelling complexes (NRCs) initiates nucleosome displacement to form narrow nucleosome-free regions at poised enhancers. The redundant histone methyltransferases myeloid/lymphoid or mixed-lineage leukaemia protein 3 (MLL3) and MLL4 deposit the active histone H3 lysine 4 monomethylation (H3K4me1) and H3K4me2 marks, whereas the histone-lysine *N*-methyltransferase EZH2 (a component of the Polycomb complex) deposits repressive H3K27me3 marks, and histone deacetylase (HDAC)-containing complexes maintain histones in a repressed, deacetylated state. RNA polymerase II (Pol II) is either absent or found at low levels at poised enhancers. **b** | In response to various cues, signal-dependent transcription factors (SDTFs) associate with recognition motifs in close association with LDTFs, which results in additional nucleosome displacement, as observed by widening of the DNase I-hypersensitive sites. SDTFs recruit co-activator complexes containing histone demethylase (HDM) complexes that remove H3K27me3 marks, histone acetyltransferases (HATs) that deposit H3K27 acetylation (H3K27ac) marks, and the Mediator complex (MED). The transformation to elongating Pol II results in bidirectional transcription — a hallmark of active enhancers — and the generation of enhancer RNAs (eRNAs), which is closely coupled to enhancer activity.

and induced pluripotent stem cells indicated that the majority of remodelling affects a single nucleosome, and that alterations in nucleosome occupancy are more frequent at enhancers that are associated with pluripotency and differentiation[17]. Transcription factor binding leads to, and in some cases is facilitated by, the recruitment of co-regulators such as the histone acetyltransferase (HAT) p300 (REF. 18), followed by the recruitment of RNA polymerase II (Pol II) and the transcription of enhancer RNAs (eRNAs)[19,20]. Co-regulator recruitment and transcription are accompanied by the covalent modification (methylation and acetylation, among others) of histone tails in enhancer-associated nucleosomes. In organisms with DNA methylation in the context of CG dinucleotides (that is, CpG methylation), these enhancers become demethylated upon their activation, concomitant with transcription factor binding[21]. Thus, epigenetic modification patterns can be used to distinguish between different enhancer activation states[22] and have been used extensively to annotate putative enhancers in different cell types[16].

Enhancer states can broadly be classified as inactive, primed, poised or active[22]. An inactive enhancer is essentially buried in compact chromatin and is devoid of transcription factor binding and histone modifications. Primed enhancers are characterized by closely bound sequence-specific transcription factors that establish a DNase I-hypersensitive[15] and nucleosome-free[23] region of open chromatin. However, they may require additional cues to accomplish their function, which may include signal-dependent activation, the recruitment of additional transcription factors and the eventual recruitment of co-activators that lead to enhancer activation. Poised enhancers can be defined as primed enhancers that also contain repressive epigenetic chromatin marks (see below), a state that is most commonly found in ESCs. The characteristic features of poised and active enhancers are depicted in FIG. 1.

An important insight for the identification of potential enhancers was the understanding that specific histone methylation signatures mark enhancer-like regions. In particular, enhancers display enrichment of histone H3 lysine 4 monomethylation (H3K4me1) or H3K4me2 and depletion of H3K4me3 compared with promoters[3]. Although genomic regions exhibiting these features are not necessarily functional enhancers, it seems that the vast majority of regions that do function as enhancers exhibit these characteristics[3,7,24]. Specifically, primed enhancer-like regions are marked with H3K4me1 and H3K4me2 and lack histone acetylation, and enhancers marked additionally by H3K27me3, a repressive mark, are considered to be poised[24–26] (reviewed in REF. 27) (FIG. 1). Features associated with active enhancers include H3K27 acetylation (H3K27ac)[25] and the presence of actively transcribing Pol II[19]. Examples of these features in the vicinity of the T-cell acute lymphocytic leukaemia 1 (*TAL1*) locus in the genomes of seven human cell lines, evaluated by the ENCODE consortium, are illustrated in FIG. 2. Several developmental enhancers have been characterized for this locus: the −3.8-kb (upstream) and +19-kb (downstream) enhancers drive *TAL1* transcription in human umbilical vein endothelial cells (HUVECs) and haematopoietic stem and progenitor cells[28,29], whereas the +51-kb enhancer is required for *TAL1* expression in K562 erythroid cells[30]. DNase I hypersensitivity at this locus corresponds with overall transcription factor binding, and the presence of the active epigenetic marks H3K4me2 and H3K27ac is correlated with cell type-specific enhancer activity. Conversely, in cells that do not express *TAL1*, such as human ESCs and normal human epidermal keratinocytes, the +19-kb enhancer, promoter and gene body are devoid of DNase I-hypersensitive sites, and the −3.8-kb region and the gene body exhibit the repressive H3K27me3 mark.

Primed enhancers
Enhancers that have been selected by lineage-determining transcription factors and collaborating transcription factors. They are marked with characteristic histone modifications such as histone H3 lysine 4 monomethylation (H3K4me1) and H3K4me2, but do not produce enhancer RNAs.

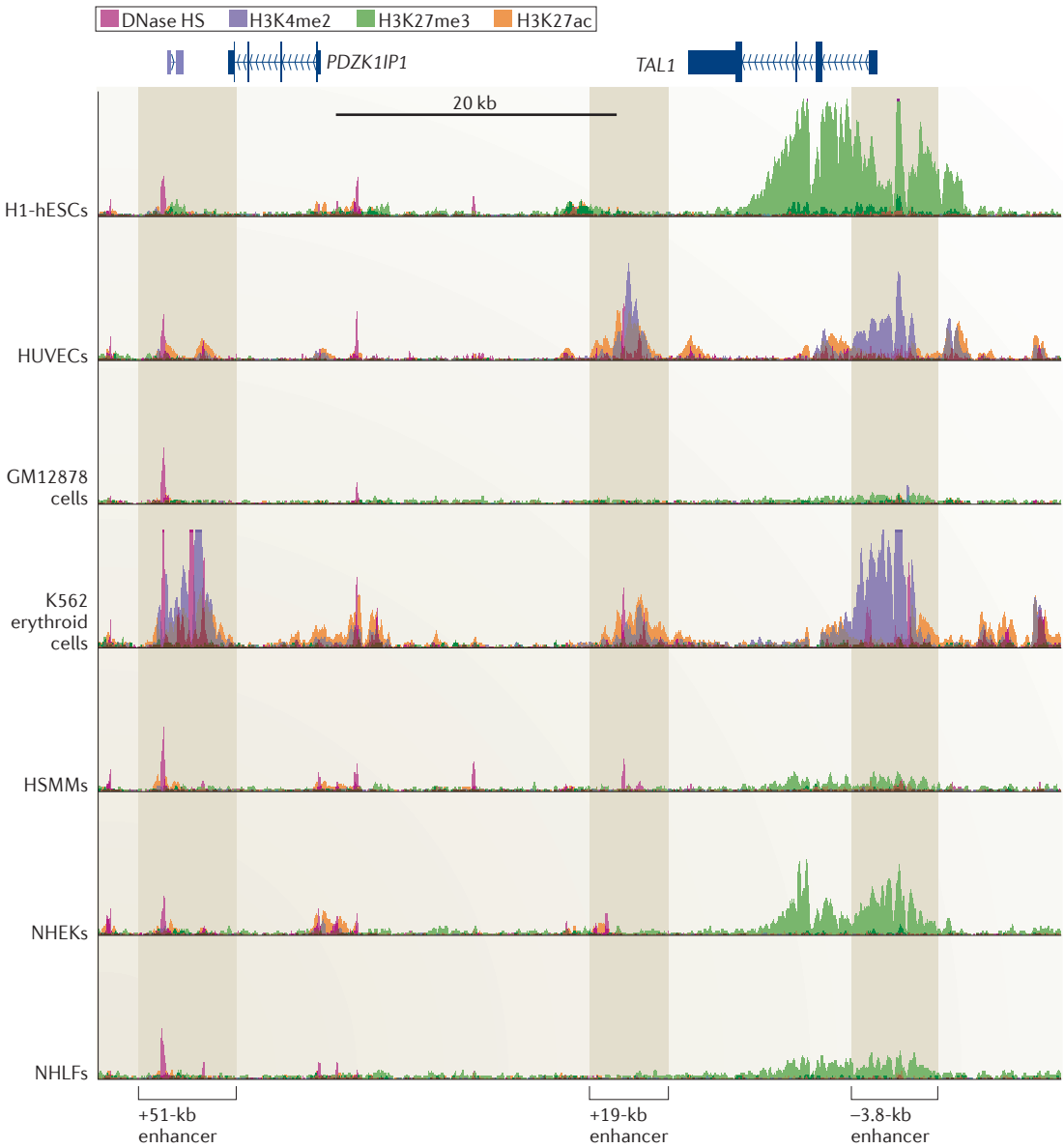

Figure 2 | **Cell type-specific enhancers are marked by specific epigenomic features and chromatin accessibility.** Genomic features of a ~60-kb region of human chromosome 1 centred around the T-cell acute lymphocytic leukaemia 1 (*TAL1*) gene are shown. These include Encyclopedia of DNA Elements (ENCODE) consortium data of DNase I-hypersensitive (DNase HS) regions, as well as chromatin immunoprecipitation followed by sequencing (ChIP–seq) data of histone H3 lysine 4 dimethylation (H3K4me2), H3K27me3 and H3K27acetylation (H3K27ac) marks in seven cell lines. Enhancers that are known to be responsible for *TAL1* transcription in endothelial cells (the –3.8-kb and +19-kb enhancers, relative to the *TAL1* promoter, in human umbilical vein endothelial cells (HUVECs)) and in erythroid cells (the +51-kb enhancer in K562 cells) exhibit cell type-specific DNase I hypersensitivity, H3K4me2 and H3K27ac. In cell types in which *TAL1* is not expressed, the promoter and gene body are devoid of DNase HS regions and histone modifications that are indicative of enhancer activation (H3K4me2 and H3K27ac), and they exhibit variable levels of the repressive H3K27me3 mark. Shaded boxes indicate cell type-restricted or cell type-specific enhancer regions. hESC, human embryonic stem cell; HSMM, human skeletal muscle myoblast; NHEK, normal human epidermal keratinocyte; NHLF, normal human lung fibroblast; *PDZK1IP1*, PDZK1-interacting protein 1.

DNase I-hypersensitive
Pertaining to genomic sites where the chromatin was made more accessible (that is, hypersensitive) to digestion by DNase I owing to the binding of regulatory proteins.

Poised enhancers
Regulatory elements that are similar to primed enhancers but that are distinguished by the presence of histone H3 lysine 27 trimethylation (H3K27me3), which must be removed to allow the transition to an active enhancer state.

Active enhancers
Enhancers that are marked with histone H3 lysine 27 acetylation (H3K27ac) marks, in addition to the marks of poised enhancers. They produce enhancer RNAs, are bound by the Mediator complex and exert regulatory functions to increase the transcription of target genes.

Enhancer selection

The vast number of potential *cis*-regulatory elements in the genome and the cell type selectivity with which they are used raise the question of how unique enhancer repertoires are selected. Many lines of evidence indicate that enhancer selection is initially driven by 'pioneer' factors, exemplified by FOXA1 (also known as HNF3α), that are able to bind to their recognition motifs within the context of compacted chromatin[31]. By opening the conformation of the chromatin and initiating the process of enhancer selection, such pioneer factors can function as key cell lineage-determining transcription factors (LDTFs) to drive lineage-specific transcription programmes. However, most sequence-specific transcription

factors, including those that function as pioneer factors, recognize fairly short DNA sequences (of approximately 6–12 bp in length), and their typical DNA recognition motifs exhibit varying levels of degeneracy. This means that most sequence-specific transcription factors have millions of potential binding sites in the mammalian genome. Nevertheless, chromatin immunoprecipitation followed by sequencing (ChIP–seq) experiments have indicated that they bind only to a small subset of all potential sites, and that a large fraction of the observed binding is associated with cell type-specific enhancers[32]. Cell type-specific binding sites often harbour motifs for additional pioneer factors, and experimental data strongly suggest that pioneer factors act in concert to jointly displace nucleosomes[33,34]. Below, we review evidence that supports a model in which pioneer factors, or LDTFs, prime cell type-specific enhancers through collaborative interactions[7,23,35–40].

The role of lineage-determining transcription factors. Experiments modulating the expression of LDTFs have demonstrated the ability of these factors to initiate the transition of enhancer elements from closed chromatin to a primed or poised state, in which transcription factors have gained access to the DNA and established nucleosome-free regions[7,10] (FIG. 1). An example is provided by the ETS domain transcription factor PU.1 (also known as SPI1), a LDTF required for the development of macrophages and B cells. PU.1 influences the establishment of distinct gene expression programmes in each cell type[41]. The vast majority of PU.1 binding sites are located >500 bp from promoters and mostly occupy different genomic locations in macrophages and B cells[7]. Macrophage-specific binding of PU.1 was observed at genomic locations that contained PU.1 binding sites in close proximity to binding sites of other macrophage LDTFs, such as the CCAAT/enhancer-binding proteins (C/EBPs) and activator protein 1 (AP1) factors. Conversely, B cell-specific binding of PU.1 was observed in close proximity to other B cell LDTFs, including motifs recognized by the transcription factor EBF1 (also known as COE1), the transcription factor E2α (E2A) and OCT factors. The corresponding motifs were generally situated <100 bp from the PU.1 motif, but most of these are not found at a close (5–20 bp) invariable distance that would be indicative of direct ternary protein–protein–DNA interactions[42]. Notably, macrophage-specific PU.1-bound regions were depleted of B-cell LDTF motifs, and vice versa, relative to adjacent genomic regions. This — together with the finding that in a given cell type, non-bound PU.1 motifs in transcriptionally inactive genomic regions are generally depleted of motifs of the LDTFs expressed in the cell[32] — suggests that LDTF motif composition may be one of the contributing factors to the formation of transcriptionally inactive and active genomic compartments (see below). Gain- and loss-of-function experiments revealed an interdependence of PU.1 with other LDTFs for effective DNA binding, which suggests that their collaborative interactions are necessary to compete with nucleosomes for binding to DNA. By considering natural genetic variation between inbred strains of mice as mutations, LDTF binding site mutations were found to impair not only binding of the respective LDTF but also that of closely bound LDTFs[43,44], which is consistent with a model in which enhancer selection is achieved through the collaborative efforts of multiple DNA-binding factors. Other examples of LDTF cooperativity in establishing specific LDTF binding patterns have been observed in developmental systems, such as zebrafish haematopoiesis[45] and *Drosophila melanogaster* embryogenesis[46].

The use of computational methods to identify binding motifs that are enriched in genomic regions marked by H3K4me1 resulted in the identification of LDTF motifs of the corresponding cell types. For example, binding sites for transcription factors that are capable of reprogramming fibroblasts into induced pluripotent stem cells are highly enriched in the H3K4me1-marked regions of the genome of ESCs[7]. Conversely, ChIP–seq studies commonly revealed that LTDFs occupy large fractions of the enhancers within the cells in which they exert lineage-determining functions[7,10,16,36,47,48]. Thus, whereas most cells express hundreds of transcription factors, the selection of a large proportion of cell type-specific regulatory elements may be driven by fairly simple combinations of LDTFs that interact with each other and with other factors. Collectively, these findings may facilitate computational efforts to predict the selection of cell type-specific enhancer elements based on the local organization of binding motifs and the combinations of expressed transcription factors.

The role of signal-dependent transcription factors. Whereas LDTF binding may be sufficient for the activation of some enhancers, additional signals will be required for other enhancers to be fully activated. Many of the cellular responses to internal and external signals are dependent on the function of widely expressed, signal-dependent transcription factors (SDTFs). Examples of SDTFs are members of the nuclear receptor and NF-κB families. These factors frequently activate common sets of genes in different cell types but can also regulate gene expression in a cell type-specific manner. ChIP–seq studies of SDTFs in different cell types found both common and cell type-specific binding sites[37,38,48]. Two types of mechanisms were suggested to account for cell type-specific binding of SDTFs. In one mechanism, such SDTF binding occurs at genomic locations that exhibit features of pre-selected enhancers[7,36,45] (FIG. 3a). In these cases, there is a hierarchical relationship between SDTFs and LDTFs, with the latter acting as pioneer factors that are responsible for the initial enhancer selection through interactions with additional collaborating transcription factors (CTFs). In many cases, loss of function of the LDTF results in a failure of both the LDTF and the SDTF to bind to the enhancer, but loss of function of SDTFs does not result in a failure of LDTF binding[7,48–51]. Alternatively, SDTFs could contribute directly to latent or *de novo* enhancer selection[37,50,52] (FIG. 3b). This has been shown to involve collaborative interactions with LDTFs which, owing to their restricted cell type-specific expression patterns, impose cell type-specific enhancer selection at genomic locations that have the appropriate combination of motifs. Although the

Latent or *de novo* enhancer
An inactive enhancer that requires the binding of a combination of transcription factors, including signal-dependent transcription factors, for selection.

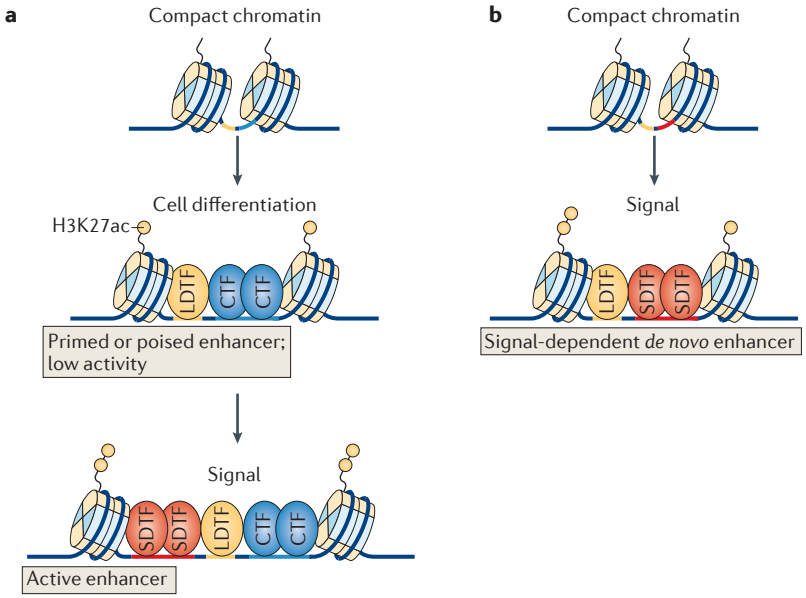

Figure 3 | **Cell type-specific enhancer selection and activation.** **a** | Collaborative interactions between lineage-determining transcription factors (LDTFs) and collaborating transcription factors (CTFs) select enhancers for binding and activation by signal-dependent transcription factors (SDTFs). Prior to signal-dependent activation, such regions may be poised enhancers or exhibit basal enhancer activity (that is, they are pre-existing enhancers) that is further induced by the binding of a SDTF. The resulting transcription is cell type-specific because the enhancers are selected by the cell type-specific LDTFs. **b** | SDTFs can direct the selection of latent or *de novo* enhancers. In these cases, the SDTF functions as an essential CTF to LDTFs to enable concurrent binding of all factors involved. The transcriptional output is cell type-specific because of the requirement for cell type-specific LDTFs for enhancer priming. H3K27ac, histone H3 lysine 27 acetylation.

mechanisms underlying collaborative DNA binding by transcription factors remain poorly understood, studies of the glucocorticoid receptor suggest that transcription factor binding can be highly dynamic, and that even two factors that interact with the same recognition motif in the same cell can facilitate each other's binding through a proposed assisted loading mechanism[53].

The extent to which SDTFs operate on poised enhancers or participate in *de novo* enhancer selection seems to vary depending on the factor, cell type and signal in question. Forkhead box protein P3 (FOXP3), a SDTF required for the acquisition of the T_H2 phenotype of $CD4^+$ T cells, was found to bind almost exclusively to poised enhancers upon their activation[36]. By contrast, the receptor for the steroid hormone ecdysone, a member of the nuclear receptor family that mediates transcriptional responses to ecdysone in insects, primarily binds to newly selected enhancer elements in combination with cell type-specific transcription factors[37]. Both mechanisms of enhancer selection (FIG. 3) can occur simultaneously in the same cell type. For example, following macrophage activation by lipopolysaccharides, approximately 90% of the binding of the p65 subunit of NF-κB occurs at enhancers that are already primed, whereas the remainder is associated with the *de novo* selection of latent enhancers in collaboration with LDTFs, such as PU.1 and C/EBPα[43,50]. Studies of different tissue macrophage populations demonstrate the importance of the environment for maintaining expression of cell type-specific transcription factors that in turn activate cell-specific enhancer programmes[54,55]. These programmes arise from a combination of *de novo* enhancer selection by collaborative interactions involving cell-restricted transcription factors, as well as from the activation of poised enhancers by cell-restricted transcription factors[54]. Of note, the histone methylation signature of latent enhancers persists after the cessation of cell stimulation and is associated with more-rapid and more-diverse transcriptional responses to subsequent stimulation[52]. These observations provide evidence that the writing of the H3K4me1 signature in enhancers provides a molecular 'memory' of prior activation.

Enhancer activation

Although transcription factor binding is a requirement for enhancer activity, not all promoter-distal transcription factor binding sites seem to function as enhancers on the basis of a lack of H3K4me1 and H3K4me2 marks, and not all regions of the genome enriched in H3K4me1 and H3K4me2 exhibit marks of active enhancers, such as H3K27ac. This raises the question of what determines whether transcription factor binding will result in an active enhancer. Many different enhancer states can be defined based on particular combinations of histone post-translational modifications[22] (FIG. 1), which are deposited by transcription co-regulators that are recruited to enhancers and promoters by transcription factors. Transcription co-regulators include histone methyltransferases, such as the myeloid/lymphoid or mixed-lineage leukaemia (MLL) proteins[56]; HATs such as p300 and CREB-binding protein (CBP)[57]; histone deacetylases, which are components of co-repressors such as nuclear receptor corepressors (NCORs) and SMRT (REF. 5); and chromatin remodellers such as the transcription activator BRG1 complex or the BRM-associated factor (BAF) complex (also known as the mammalian SWI/SNF complexes)[58,59] and the Mediator complex[60]. Recruitment of co-regulators to a given enhancer is more frequent when more transcription factors are co-bound to that enhancer[7,61]. Co-regulators are large proteins with multiple distinct interaction sites for transcription factors[18,62,63] and probably act as both facilitators and integrators of transcription factor binding and intracellular signals at enhancers, which is similar to their known roles at promoters[64].

Enhancer transcription. The epigenetic marks deposited by co-regulator complexes act as binding sites for chromatin 'readers' such as TFIID[65] and the bromodomain-containing protein 4 (BRD4)–positive transcription elongation factor-b (P-TEFb) complex[66], which function in transcription pre-initiation complex assembly and in transcription elongation, respectively.

The presence of the transcription pre-initiation complex and elongation factors at enhancers[67,68] is in line with the finding that Pol II is found at enhancers. More than 20 years ago Pol II was observed to generate non-coding RNAs at locus control regions[69], but it was only recently appreciated that mammalian enhancers are broadly transcribed and generate eRNAs[19,20,70–72]. Pol II recruitment to enhancers and signalling-dependent changes in

Mediator complex
A protein complex with important roles in transcription and the 3D organization of the genome. It integrates various intracellular signals to affect the formation of the pre-initiation complex, transcription initiation and transcription elongation.

Locus control regions
Regions that confer tissue-specific expression to linked transgenes irrespective of the transgene integration site in the genome. These regions display characteristics of both enhancers and insulators.

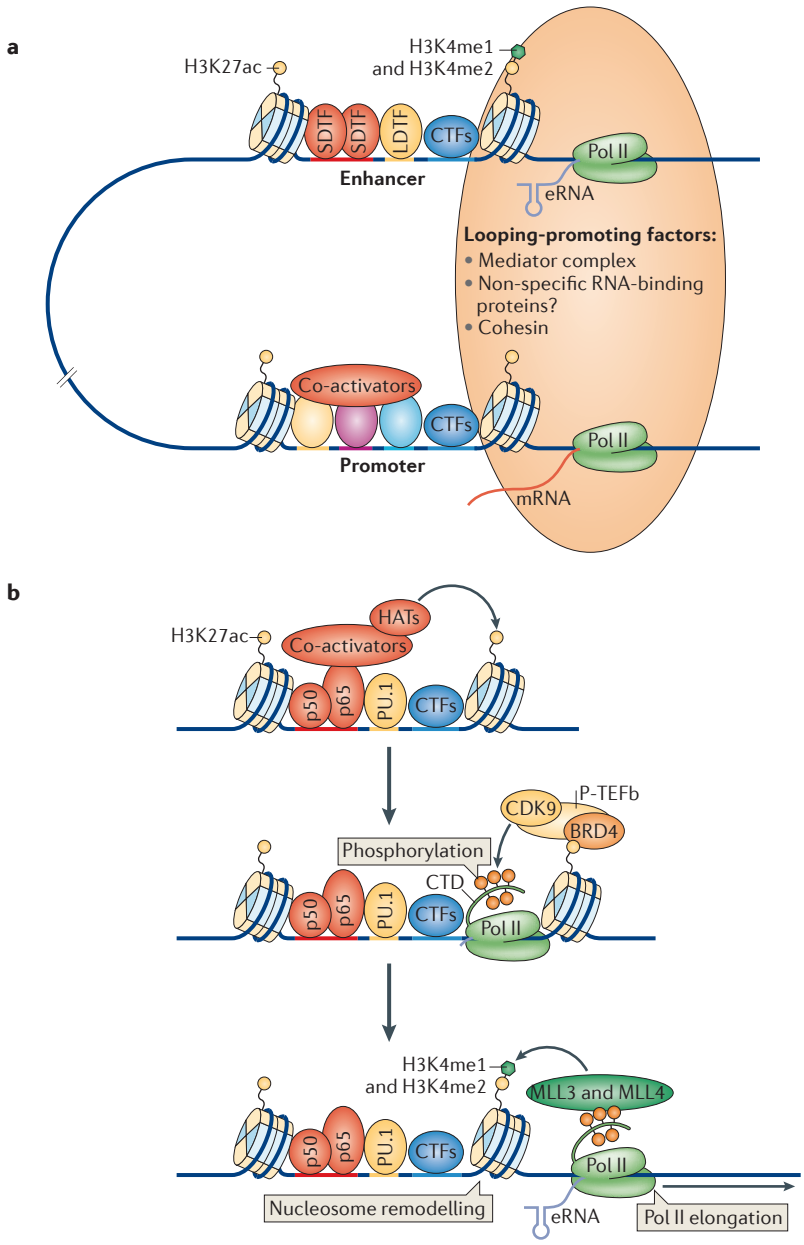

Figure 4 | **Enhancer activation and function.** **a** | Interactions between enhancers and promoters involve structural connections (orange oval) that include cohesin and the Mediator complex to promote formation of the pre-initiation complex, to initiate transcription and/or to overcome RNA polymerase II (Pol II) pausing. A potential role of enhancer RNAs (eRNAs) could be to promote transcription by facilitating chromatin looping, possibly by mediating interactions with cohesins. Another potential role of eRNAs could be to mediate interactions with protein complexes that are required for transcription elongation, such as the Mediator complex. **b** | In activated macrophages, the nuclear factor-κB (NF-κB) subunits p50 and p65, which are signal-dependent transcription factors (SDTFs), and the lineage-determining transcription factor (LDTF) PU.1 collaboratively select *de novo* enhancers. The subsequent recruitment of histone acetyltransferases (HATs) results in histone acetylation, a mark that is subsequently bound by the bromodomain-containing protein 4 (BRD4)–positive transcription elongation factor-b (P-TEFb) complex, which allows its cyclin-dependent kinase 9 (CDK9) component to phosphorylate the carboxy-terminal domain (CTD) of Pol II. The phosphorylated CTD acts as a docking site for the myeloid/lymphoid or mixed-lineage leukaemia protein 3 (MLL3) and MLL4 histone H3 lysine 4 (H3K4) methyltransferases, which are proposed to deposit H3K4 monomethylation (H3K4me1) and H3K4me2 marks during successive rounds of Pol II elongation. CTF, collaborating transcription factor; H3K27ac, H3K27 acetylation.

eRNA expression are highly correlated with changes in the expression of nearby genes, which suggests a functional link between eRNA and gene expression[50,73–75]. The distinguishing features of eRNAs are that most are short (<1 kb), are not subjected to polyadenylation or splicing[19,20] and are rapidly degraded by the exosome[71]. Similarly to what has been shown for short promoter antisense transcripts[76], these characteristics are probably caused by the lack of 5′ splice donors proximal to eRNA TSSs[71,72], which are prerequisites for splicing and promote elongation[77], packaging into messenger ribonucleoprotein particles (mRNPs), polyadenylation and nuclear export[78], all of which contribute to the stability of transcripts. As a side note, the fact that enhancers resemble promoters in almost every aspect, apart from the lack of proximal splice donors[71] and H3K4me3 marks[79], suggests that stable mRNAs or lincRNAs could be created by simply introducing a splice donor downstream of an eRNA TSS[72]. This would be consistent with the ability of intronic enhancers to act as alternative promoters[80] and with the fact that 42% of all lincRNAs have only two exons (that is, a single splice donor downstream of a promoter), compared with merely 6% of coding genes[81].

The occurrence of global enhancer transcription in mammalian cells raises the question of its functional importance. Recent studies provide evidence that eRNAs contribute to local enhancer activity, potentially by facilitating enhancer–promoter interactions through chromatin looping, recruitment of cofactors such as the Mediator complex (reviewed in REF. 66) (FIG. 4a) and the release of the negative elongation factor complex[82]. So far, there is limited evidence for specific sequence features of eRNAs that could be necessary for their function, and not all eRNAs seem to contribute to enhancer function. To date, little attention has been directed at the possibility that the process of enhancer transcription itself (which is independent of the eRNA product) could influence enhancer activity. Pol II is a powerful nucleosome-remodelling machine[83], and transcription initiated from an enhancer sequence may contribute to maintaining an open chromatin configuration that enables access of sequence-specific transcription factors. In addition, enhancer transcription may have an important role in contributing to the deposition of H3K4me1 and H3K4me2 marks at enhancers (FIG. 4b). Genetic studies indicate that the *D. melanogaster* H3K4 methyltransferase Trithorax-related (Trr) and its mammalian homologues MLL3 and MLL4 play important parts in the deposition of these marks[84,85], but the mechanisms that recruit these enzymes and that determine the overall distribution of histone methylation remain poorly understood. Studies of newly selected or *de novo* enhancers in activated macrophages provided evidence that the methylation of H3K4, but not the acetylation of H3K27, required enhancer transcription and the presence of MLL3 and MLL4 (REF. 50).

A model of enhancer activation based on time-resolved studies of transcription factor binding, eRNA transcription, H3K4 methylation and H3K27 acetylation at *de novo* enhancers, and on results of gain- and loss-of-function experiments[50], is illustrated in FIG. 4b. Signal-dependent activation of NF-κB (which is composed of subunits

p50 and p65) results in its collaborative binding to the enhancer with PU.1 and the recruitment of co-activator complexes that contain HATs. These events result in nucleosome remodelling, histone acetylation and the recruitment of Pol II. The conversion of Pol II from a paused form to an elongating form involves the BRD4–P-TEFb complex, which is recruited to at least some sites of transcription initiation by interactions between BRD4 and acetylated histone H4. Cyclin-dependent kinase 9 (CDK9), a component of P-TEFb, phosphorylates the carboxy-terminal domain (CTD) of Pol II, which provides docking sites for MLL3 and MLL4. MLL3 and MLL4 progressively methylate H3K4 during successive rounds of transcription elongation. Consistent with this model is the distribution of H3K4me1 and H3K4me2, which was found to correlate with the extent of enhancer transcription and to be dependent on transcription elongation[50]. The generality of this model with respect to the mechanisms by which H3K4 methylation marks are established at other classes of enhancers, such as those that are selected during cellular differentiation, remains to be determined. For example, in contrast to the activation of *de novo* enhancers in the context of extracellular signalling responses, studies of the distribution of H3K4me1 and H3K4me2 at cell type-specific enhancers selected during muscle and adipocyte differentiation suggested that MLL complexes can interact directly with LDTFs such as C/EBPβ and myoblast determination protein 1 (MYOD1) at cell type-specific enhancers, where MLL3 and MLL4 are also required for acetylation of H3K27 and for recruitment of the Mediator complex and Pol II[85].

The function of H3K4me1 and H3K4me2 marks remains an open question. As they are known to recruit histone-remodelling complexes[86], they could conceivably contribute to keeping previously bound and modified enhancers open and accessible, which would help to explain the observation that previously activated latent enhancers are more rapidly re-activated by subsequent stimuli[52].

Enhancer function

Promoter activation requires that many components of the transcriptional machinery come together in order to assemble the pre-initiation complex, initiate transcription, overcome Pol II pausing and eventually lead to productive transcription elongation. Through looping of the intervening DNA, enhancers are brought into close proximity of promoters and are thought to affect any or all of the aforementioned processes by increasing the local concentrations of the factors involved[87] (FIG. 4a). These factors include co-activator complexes such as the Mediator complex, which increases the loading of transcription factors onto promoters and enhancers[60]; scaffold proteins, such as cohesin, that mediate stable and often cell type-specific promoter–enhancer interactions[60,88]; and factors that are involved in releasing paused Pol II and in the initiation of elongation, such as BRD4 (REF. 89). A major challenge in deciphering cell type-specific enhancer functions is connecting active enhancers to their target genes *in vivo*.

Super-enhancers. On the basis of their epigenetic features and depending on the experimental methods used to define enhancers, ~10,000–50,000 putative enhancers can be identified in a given cell type[13,16,90], which implies that there are more enhancers than expressed genes. Along the linear DNA molecule, enhancers are located non-uniformly with respect to genes, such that some genes are located in enhancer-rich regions of the genome, whereas others have few or no enhancers in their vicinity. Although a single enhancer is sufficient to activate the expression of a nearby gene[37], high levels of cell type-specific and/or signal-dependent gene expression are most frequently observed for genes located in enhancer-rich regions of the genome, which is exemplified by the relationship between enhancer-rich locus control regions and the expression of the globin genes in erythroid cells[69]. Such enhancer-dense regions have recently been termed super-enhancers[91–93].

Super-enhancers were initially defined as large (tens of kilobases in length) genomic loci with an unusually high density of enhancer-associated marks, such as binding of the Mediator complex, relative to most other genomic loci[91,92]. These regions can also be defined by high-density[91] and/or extended (>3 kb)[94] depositions of the histone mark H3K27ac. Using differences in the density of Mediator complex-binding sites or of H3K27ac marks to distinguish super-enhancers from regular enhancers, most cell types are found to have 300–500 super-enhancers[91]. A substantial fraction of super-enhancers and nearby genes are cell type-specific, and the gene sets that are associated with super-enhancers in a given cell type are highly enriched for the biological processes that define the identities of the cell types[91,94]. For example, many of the genes encoding factors required for pluripotency and self-renewal of ESCs are located near ESC-specific super-enhancers[91]. Consistent with their tissue specificity, super-enhancers that are active in certain cell types are enriched for disease-associated alleles relevant to that cell type[91,94]. Not surprisingly, the individual enhancers of cell type-specific super-enhancers are enriched for binding sites of the corresponding LDTFs[92]. Collectively, the specific set of super-enhancers within a particular cell type may provide a means of simplifying the problem of defining what are the quantitatively most important transcriptional programmes required for establishing cell identity, and the problem of identifying disease-relevant non-coding genetic variation.

3D chromatin interactions. In the nucleus, the genome is organized and partitioned into functional compartments in 3D space[95], and considerable effort is being directed at understanding enhancer function in the context of 3D chromatin interactions. One strategy is to identify the long-range looping interactions that involve enhancer elements using a variety of chromosome conformation capture (3C)-based techniques[96]. Genome-wide applications of these techniques to define the chromatin interactomes of human and mouse cells confirmed that the genome is divided into active and inactive compartments[96]. These are further organized into submegabase-sized topologically associated domains (TADs) that

Exosome
A protein complex involved in the quality control, maturation and degradation of various RNA transcripts, both in the nucleus and cytoplasm.

Chromosome conformation capture
(3C). A method to probe the higher-order structure of the genome by capturing and sequencing DNA sites that are spatially close to each other in the nucleus.

Topologically associated domains
(TADs). Largely self-interacting genomic domains of submegabase size that are further organized into multi-megabase-sized structures called nuclear compartments. Genes within TADs are co-regulated, and their expression patterns are highly correlated.

correlate with regions of the genome that constrain the spread of heterochromatin and are relatively conserved across cell types[97]. Although the genome-wide resolution of such studies remains limited, the resulting chromatin connectivity maps suggest that only approximately 7% of the looping interactions are made between adjacent genes, which indicates that linear genomic adjacency is not necessarily a good predictor of long-range interactions[98]. In addition, promoters and distal enhancer elements are frequently engaged in multiple long-range interactions and form active chromatin hubs[98,99] (FIG. 5). Whereas super-enhancers are identified along the linear DNA sequence by means of their high density of typical epigenomic features, it is clear that the enhancers within a super-enhancer form 3D interactions that are a feature of the folded genome in the nucleus[93] (FIG. 5). Interestingly, studies of tumour necrosis factor-α (TNFα)-responsive enhancers in human fibroblasts indicated that they are already in contact with their target promoters before the activation of TNFα signalling[100], which suggests that enhancer–promoter interactomes are already set up during development. This is consistent with data from *D. melanogaster*, which show that only 6% of spatial genome interactions change during early embryonic development[101]. It is not yet known when higher-order chromatin interactions are established during development, but it is likely to coincide with the occurrence of gap phases following the mid-blastula transition, which is accompanied by the establishment of a non-random nuclear chromatin conformation and the transcriptional activation of chromatin domains[102].

It is unclear how the 3D organization of the genome is determined; however, cohesin and the Mediator complex[60] — which are scaffold proteins of the replication machinery and the transcription machinery, respectively — are known to be involved in the formation of higher-order chromatin structures. As cohesin seems to be recruited to enhancers through clusters of LDTFs[103], it is likely that both protein–protein interactions and the genomic sequence shape 3D genomic conformations, although this hypothesis still awaits experimental confirmation.

Given that the conformation of the genome seems to be mostly fixed across developmental stages[101], individual cells[104], cell types[97] and signalling states[100], it is tempting to speculate how enhancers work in the 3D space: promoters are known to also function as transcriptional enhancers with regard to the activation of promoters in their proximity[105], and enhancers have sequence features that are identical to those found in promoters[71,72]. Both are juxtaposed within TADs as part of linear super-enhancers[92] and are being brought into proximity by higher-order chromatin conformations[74], which leads to the co-regulation of promoters and enhancers within a domain[74,106]. Knocking out enhancers within a TAD shows that the loss of an enhancer often only leads to a graded reduction in expression[107,108,74] and in developmental dysregulation[109] of the associated gene, which suggests that, at least in some cases, enhancers work in an additive manner. The distribution of gene regulation among a multitude of enhancers — some of which are located linearly within or beyond neighbouring genes (in their 'shadow', hence the term 'shadow enhancers' (REF. 110)) but are close in 3D space — is thought to increase robustness of the regulatory system to mutations[111]. Whereas the higher-order chromatin structure of genomic regions >1 Mb is invariant to a large extent, single loci can transition between inert and active chromosomal states, depending on their activation status, which leads to stable repression or to a state poised for transcription, respectively[112]. By contrast, inside TADs within these large-scale compartments, the chromatin structure of regions <100 kb does differ in a cell type-specific manner[100], which implies that different regulatory regions within TADs can be dynamically juxtaposed in a stimulus-specific manner. In this way, genome topology could contribute to cell type-specific transcription programmes, which means that mapping the genomic topology and elucidating the mechanisms that govern the 3D structure of the genome will be important steps towards understanding how the genome functions.

Chromatin hubs
Nuclear domains comprised of regulatory DNA elements (locus control regions, enhancers and promoters) and genes that enable correct gene expression. The smallest unit of a hub could be a topologically associated domain, and the largest could comprise an entire nuclear compartment.

Conclusions and perspective

Although initially described >30 years ago, we still do not have a clear understanding of the mechanisms by which enhancers regulate gene expression. However, the development of a plethora of methods for genome-wide mapping of diverse enhancer features, their functional relationship with promoters and their ultimate transcriptional outputs has resulted in a number of striking and unexpected discoveries, ranging from the identification of the great number of enhancers in metazoan genomes to the widespread production of eRNAs. The observation that >80% of disease-associated alleles identified by genome-wide association studies are found in non-coding regions of the genome[113] implies that they have yet unappreciated regulatory functions. Consistent with this finding, several studies have demonstrated an enrichment of disease-associated loci in cell type-specific regulatory regions, including in super-enhancers, of the corresponding disease-relevant cell types[91,114–117], and a number of studies are beginning to document the direct effects of common variation in enhancer elements on enhancer states[118–120], gene expression[117,121,122] and disease[123–127].

Beyond the simple annotation of regulatory regions in the genome, it is important to understand how cells select the full complement of enhancers that are required for maintaining their identities and functions. In essence, we would like to be able to read the genomic template and predict from the combination of active transcription factors the enhancers that will be functional in a cell type-specific manner. The principle of collaborative transcription factor interactions at closely spaced DNA recognition motifs provides a starting point for predicting genome-wide patterns of transcription factor binding that are required for enhancer selection. These predictions can be validated by mutating binding sites or by taking advantage of naturally occurring genetic variation. However, transcription factor binding maps are insufficient for predicting enhancer activity. The discovery that enhancer transcription is highly correlated with nearby gene expression is likely to be an important clue in understanding how enhancers function. The evidence

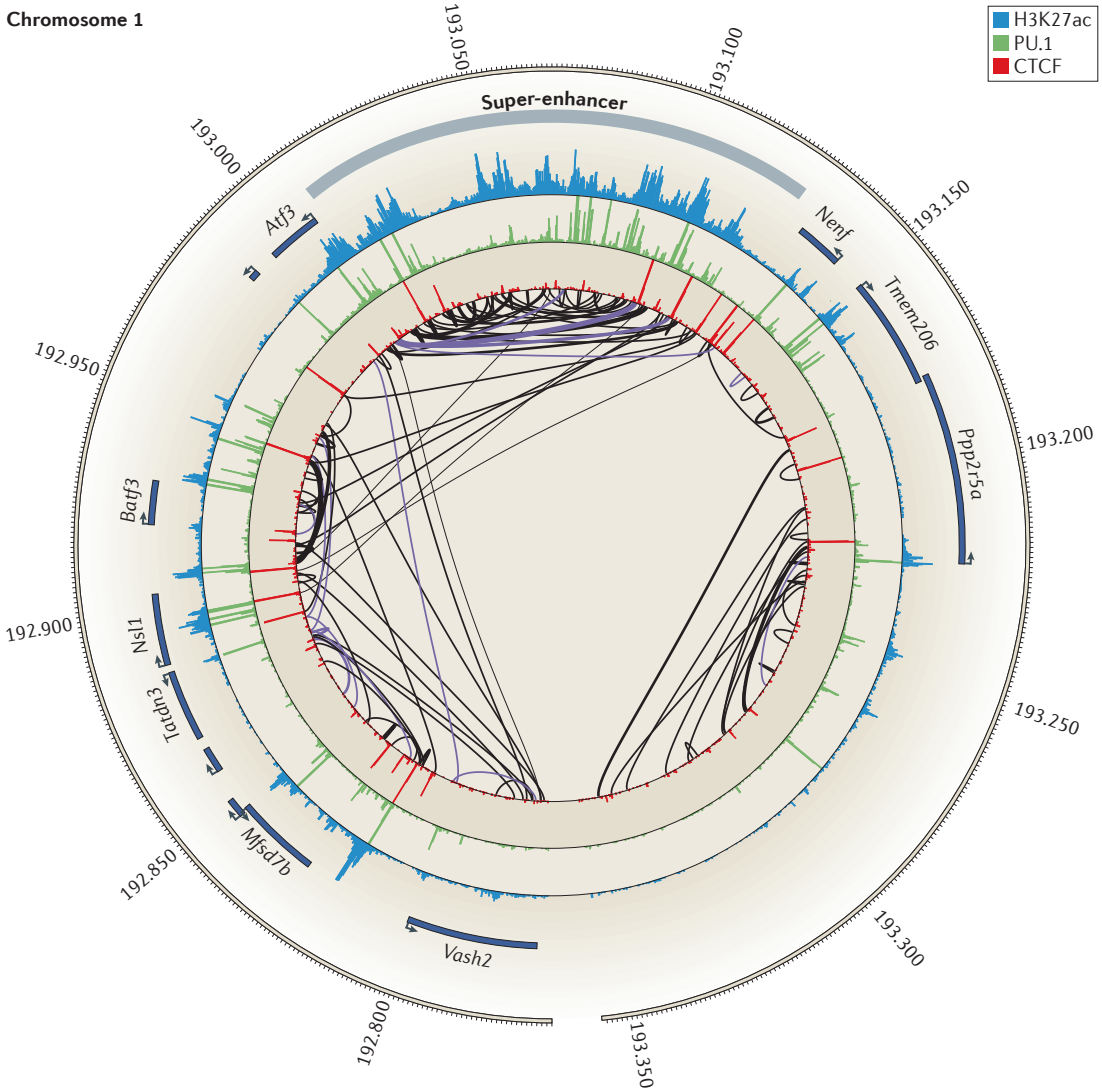

Figure 5 | **The linear and 3D organization of enhancers in the nucleus.** The outer circle represents the linear coordinates of a region of mouse chromosome 1 (mm9 assembly) surrounding the activating transcription factor 3 (*Atf3*) gene in C57BL/6J mouse macrophages. The locations of individual genes are indicated by gene names and blue bars. The three successive concentric inner circles depict chromatin immunoprecipitation followed by sequencing (ChIP–seq) data of histone H3 lysine 27 acetylation (H3K27ac), the transcription factor PU.1 and the transcription repressor CCCTC-binding factor (CTCF), which is enriched at the boundaries of topologically associated domains (TADs). A region of high density of H3K27ac in the vicinity of the *Atf3* gene is designated as a super-enhancer. Purple and black lines in the centre of the circle indicate physical contacts involving promoters and other genomic regions, respectively, as determined by significant genome-wide chromatin connectivity measurements using tethered conformation capture. This locus demonstrates the multitude of connections between the individual enhancers that constitute the *Atf3* super-enhancer, which essentially forms its own TAD, as well as the longer-range enhancer–enhancer and enhancer–promoter interactions outside the TAD. *Batf3*; basic leucine zipper transcription factor, ATF-like 3; *Mfsd7b*, major facilitator superfamily domain-containing 7B; *Nenf*, neudesin neurotrophic factor; *Nsl1*, NSL1, MIND kinetochore complex component, homologue; *Ppp2r5a*, protein phosphatase 2, regulatory subunit B alpha; *Tatdn3*, TatD DNase domain-containing 3; *Tmem206*, transmembrane protein 206; *Vash2*, vasohibin 2.

that eRNAs contribute to the activities of at least some enhancers provides momentum for determining their mechanisms of action. In addition, the importance of enhancer transcription itself in maintaining enhancer accessibility and contributing to enhancer-related H3K4 methylation requires further study, as the functional roles of H3K4 methylation, beyond providing a memory of prior enhancer activation, remain obscure[52].

Defining functional enhancer–promoter interactions remains an important goal. Despite being informative, chromatin connectivity maps do not directly relate chromatin interactions to the regulation of gene expression. Definitive evidence that a specific enhancer-like region exerts a transcriptional regulatory function requires the study of mutational effects on that region and, encouragingly, site-specific mutagenesis should

be greatly facilitated by recently developed genome editing methods[128]. Such tools will enable us to systematically delete enhancer elements and modify enhancer sequences to evaluate chromatin connectivity and gene expression. As a complementary approach, recent studies have demonstrated the use of natural genetic variation as a tool to study the relationships between transcription factor binding, enhancer selection and the regulation of gene expression[43,54]. Improving the understanding of the mechanisms underlying the selection and function of enhancers is likely to not only enable prediction of the consequences of genetic variation on gene expression and phenotype, but also provide approaches to directly alter enhancer function for therapeutic purposes.

1. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308 (1981).
2. Lettice, L. A. *et al.* 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 (2003).
3. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genet.* **39**, 311–318 (2007).
4. Carroll, J. S. *et al.* Genome-wide analysis of estrogen receptor binding sites. *Nature Genet.* **38**, 1289–1297 (2006).
5. Barish, G. D. *et al.* Bcl-6 and NF-κB cistromes mediate opposing regulation of the innate immune response. *Genes Dev.* **24**, 2760–2765 (2010).
6. John, S. *et al.* Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nature Genet.* **43**, 264–268 (2011).
7. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
This paper shows that the LDTF PU.1 binds to different loci in different cell types. The loci are marked by H3K4me1, act as beacons for SDTFs binding and drive cell type-specific cellular responses.
8. Lefterova, M. I. *et al.* Cell-specific determinants of peroxisome proliferator-activated receptor γ function in adipocytes and macrophages. *Mol. Cell. Biol.* **30**, 2078–2089 (2010).
9. Nielsen, R. *et al.* Genome-wide profiling of PPARγ:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev.* **22**, 2953–2967 (2008).
10. Ghisletti, S. *et al.* Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32**, 317–328 (2010).
This paper shows that PU.1 is an LDTF necessary for priming signal-dependent enhancers in macrophages, which defines their transcriptional response to inflammatory stimuli.
11. Pennacchio, L. A. *et al. In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499–502 (2006).
12. Woolfe, A. *et al.* Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol.* **3**, e7 (2005).
13. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
14. Visel, A. *et al.* ChIP–seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
15. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).
16. Bernstein, B. E. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
This paper summarizes the work of the ENCODE consortium to annotate functional DNA elements in the human genome.
17. West, J. A. *et al.* Nucleosomal occupancy changes locally over key regulatory regions during cell differentiation and reprogramming. *Nature Commun.* **5**, 4719 (2014).
18. Chan, H. M. & La Thangue, N. B. p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell Sci.* **114**, 2363–2373 (2001).
19. De Santa, F. *et al.* A large fraction of extragenic RNA Pol II transcription sites overlap enhancers. *PLoS Biol.* **8**, e1000384 (2010).
20. Kim, T.-K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
References 19 and 20 report the existence of widespread transcription at enhancers, which correlates with the transcription of neighbouring genes.
21. Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
22. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature Biotech.* **28**, 817–825 (2010).
23. He, H. H. *et al.* Nucleosome dynamics define transcriptional enhancers. *Nature Genet.* **42**, 343–347 (2010).
24. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).
25. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl Acad. Sci. USA* **107**, 21931–21936 (2010).
26. Zentner, G. E., Tesar, P. J. & Scacheri, P. C. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* **21**, 1273–1283 (2011).
27. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Mol. Cell* **49**, 825–837 (2013).
28. Gottgens, B. *et al.* The *scl* +18/19 stem cell enhancer is not required for hematopoiesis: identification of a 5′ bifunctional hematopoietic-endothelial enhancer bound by Fli-1 and Elf-1. *Mol. Cell. Biol.* **24**, 1870–1883 (2004).
29. Sanchez, M. *et al.* An *SCL* 3′ enhancer targets developing endothelium together with embryonic and adult haematopoietic progenitors. *Development* **126**, 3891–3904 (1999).
30. Delabesse, E. *et al.* Transcriptional regulation of the *SCL* locus: identification of an enhancer that targets the primitive erythroid lineage *in vivo*. *Mol. Cell. Biol.* **25**, 5215–5225 (2005).
31. Zaret, K. S. *et al.* Pioneer factors, genetic competence, and inductive signaling: programming liver and pancreas progenitors from the endoderm. *Cold Spring Harb. Symp. Quant. Biol.* **73**, 119–126 (2008).
32. Pham, T. H. *et al.* Mechanisms of *in vivo* binding site selection of the hematopoietic master transcription factor PU.1. *Nucleic Acids Res.* **41**, 6391–6402 (2013).
33. Adams, C. C. & Workman, J. L. Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Mol. Cell. Biol.* **15**, 1405–1421 (1995).
34. Boyes, J. & Felsenfeld, G. Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. *EMBO J.* **15**, 2496–2507 (1996).
35. Hoffman, B. G. *et al.* Locus co-occupancy, nucleosome positioning, and H3K4me1 regulate the functionality of FOXA2-, HNF4A-, and PDX1-bound loci in islets and liver. *Genome Res.* **20**, 1037–1051 (2010).
36. Samstein, R. M. *et al.* Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* **151**, 153–166 (2012).
37. Shlyueva, D. *et al.* Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin. *Mol. Cell* **54**, 180–192 (2014).
38. Vahedi, G. *et al.* STATs shape the active enhancer landscape of T cell populations. *Cell* **151**, 981–993 (2012).
39. Xu, J. *et al.* Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis. *Dev. Cell.* **23**, 796–811 (2012).
40. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
41. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573–1577 (1994).
42. Kazemian, M., Pham, H., Wolfe, S. A., Brodsky, M. H. & Sinha, S. Widespread evidence of cooperative DNA binding by transcription factors in *Drosophila* development. *Nucleic Acids Res.* **41**, 8237–8252 (2013).
43. Heinz, S. *et al.* Effect of natural genetic variation on enhancer selection and function. *Nature* **503**, 487–492 (2013).
44. Stefflova, K. *et al.* Cooperativity and rapid evolution of cobound transcription factors in closely related mammals. *Cell* **154**, 530–540 (2013).
45. Trompouki, E. *et al.* Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. *Cell* **147**, 577–589 (2011).
46. Yanez-Cuna, J. O., Dinh, H. Q., Kvon, E. Z., Shlyueva, D. & Stark, A. Uncovering *cis*-regulatory sequence requirements for context-specific transcription factor binding. *Genome Res.* **22**, 2018–2030 (2012).
47. Mercer, E. M. *et al.* Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity* **35**, 413–425 (2011).
48. Mullen, A. C. *et al.* Master transcription factors determine cell-type-specific responses to TGF-β signaling. *Cell* **147**, 565–576 (2011).
References 36 and 48 show that SDTFs bind to open chromatin regions that are defined by combinations of LDTFs.
49. Carroll, J. S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
50. Kaikkonen, M. U. *et al.* Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell* **51**, 310–325 (2013).
This paper shows that the SDTF NF-κB contributes to the *de novo* priming (and activation) of enhancers in conjunction with LDTFs, and that transcription elongation is required for the methylation of the flanking H3K4 by MLL3 and MLL4.
51. Sullivan, A. L. *et al.* Serum response factor utilizes distinct promoter- and enhancer-based mechanisms to regulate cytoskeletal gene expression in macrophages. *Mol. Cell. Biol.* **31**, 861–875 (2011).
52. Ostuni, R. *et al.* Latent enhancers activated by stimulation in differentiated cells. *Cell* **152**, 157–171 (2013).
This paper describes signal-dependent selection of latent enhancers in macrophages and their persistence after signal termination as a molecular memory of prior activation.
53. Voss, T. C. *et al.* Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. *Cell* **146**, 544–554 (2011).
54. Gosselin, D. *et al.* Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. *Cell* **159**, 1327–1340 (2014).
55. Lavin, Y. *et al.* Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* **159**, 1312–1326 (2014).
References 54 and 55 demonstrate how specific tissue environments differentially influence the selection and function of enhancers in macrophage populations.

56. Herz, H. M., Hu, D. & Shilatifard, A. Enhancer malfunction in cancer. *Mol. Cell* **53**, 859–866 (2014).
57. Wang, Z. *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019–1031 (2009).
58. Euskirchen, G. M. *et al.* Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. *PLoS Genet.* **7**, e1002008 (2011).
59. Morris, S. A. *et al.* Overlapping chromatin-remodeling systems collaborate genome wide at dynamic chromatin transitions. *Nature Struct. Mol. Biol.* **21**, 73–81 (2014).
60. Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430–435 (2010).
61. Siersbaek, R. *et al.* Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Rep.* **7**, 1443–1455 (2014).
62. Rosenfeld, M. G., Lunyak, V. V. & Glass, C. K. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev.* **20**, 1405–1428 (2006).
63. Blazek, E., Mittler, G. & Meisterernst, M. The mediator of RNA polymerase II. *Chromosoma* **113**, 399–408 (2005).
64. Malik, S. & Roeder, R. G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature Rev. Genet.* **11**, 761–772 (2010).
65. Vermeulen, M. *et al.* Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 (2007).
66. Dey, A., Chitsaz, F., Abbasi, A., Misteli, T. & Ozato, K. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc. Natl Acad. Sci. USA* **100**, 8758–8763 (2003).
67. Koch, F. *et al.* Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nature Struct. Mol. Biol.* **18**, 956–963 (2011).
68. Zhang, W. *et al.* Bromodomain-containing protein 4 (BRD4) regulates RNA polymerase II serine 2 phosphorylation in human $CD4^+$ T cells. *J. Biol. Chem.* **287**, 43137–43155 (2012).
69. Collis, P., Antoniou, M. & Grosveld, F. Definition of the minimal requirements within the human β-globin gene and the dominant control region for high level expression. *EMBO J.* **9**, 233–240 (1990).
70. Lam, M. T. *et al.* Rev–Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* **498**, 511–515 (2013).
71. Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
72. Core, L. J. *et al.* Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nature Genet.* **46**, 1311–1320 (2014).
73. Wang, D. *et al.* Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* **474**, 390–394 (2011).
74. Kieffer-Kwon, K. R. *et al.* Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* **155**, 1507–1520 (2013).
75. Bonn, S. *et al.* Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nature Genet.* **44**, 148–156 (2012).
76. Almada, A. E., Wu, X., Kriz, A. J., Burge, C. B. & Sharp, P. A. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013).
77. Fong, Y. W. & Zhou, Q. Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929–933 (2001).
78. Muller-McNicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nature Rev. Genet.* **14**, 275–287 (2013).
79. Bieberstein, N. I., Carrillo Oesterreich, F., Straube, K. & Neugebauer, K. M. First exon length controls active chromatin signatures and transcription. *Cell Rep.* **2**, 62–68 (2012).
80. Kowalczyk, M. S. *et al.* Intragenic enhancers act as alternative promoters. *Mol. Cell* **45**, 447–458 (2012).
81. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789 (2012).
82. Schaukowitch, K. *et al.* Enhancer RNA facilitates NELF release from immediate early genes. *Mol. Cell* **56**, 29–42 (2014).
83. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
84. Herz, H. M. *et al.* Enhancer-associated H3K4 monomethylation by Trithorax-related, the *Drosophila* homolog of mammalian Mll3/Mll4. *Genes Dev.* **26**, 2604–2620 (2012).
85. Lee, J. E. *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* **2**, e01503 (2013).
86. Sims, R. J. 3rd & Reinberg, D. Histone H3 Lys 4 methylation: caught in a bind? *Genes Dev.* **20**, 2779–2786 (2006).
87. Plank, J. L. & Dean, A. Enhancer function: mechanistic and genome-wide insights come together. *Mol. Cell* **55**, 5–14 (2014).
88. Schmidt, D. *et al.* A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20**, 578–588 (2010).
89. Liu, W. *et al.* Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* **155**, 1581–1595 (2013).
90. Nord, A. S. *et al.* Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* **155**, 1521–1531 (2013).
91. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
92. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
References 91 and 92 describe genomic regions near genes, which are highly enriched for marks of active enhancers and have essential roles in determining cell identity and function.
93. Dowen, J. M. *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374–387 (2014).
94. Parker, S. C. *et al.* Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl Acad. Sci. USA* **110**, 17921–17926 (2013).
95. Cremer, T. *et al.* Chromosome territories — a functional nuclear landscape. *Curr. Opin. Cell Biol.* **18**, 307–316 (2006).
96. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
97. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
98. Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109–113 (2012).
References 97 and 98 use global 3C assays to interrogate the 3D organization of functional elements within the genome.
99. de Laat, W. & Grosveld, F. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.* **11**, 447–459 (2003).
100. Jin, F. *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503**, 290–294 (2013).
101. Ghavi-Helm, Y. *et al.* Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**, 96–100 (2014).
102. Vassetzky, Y., Hair, A. & Mechali, M. Rearrangement of chromatin domains during development in *Xenopus*. *Genes Dev.* **14**, 1541–1552 (2000).
103. Yan, J. *et al.* Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* **154**, 801–813 (2013).
104. Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2013).
105. Li, G. *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84–98 (2012).
This study shows that promoters can function as enhancers.
106. Li, W. *et al.* Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013).
107. Bender, M. A. *et al.* The hypersensitive sites of the murine β-globin locus control region act independently to affect nuclear localization and transcriptional elongation. *Blood* **119**, 3820–3827 (2012).
108. Sur, I. K. *et al.* Mice lacking a *Myc* enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. *Science* **338**, 1360–1363 (2012).
109. Rosenbauer, F. *et al.* Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nature Genet.* **36**, 624–630 (2004).
110. Hong, J. W., Hendrix, D. A. & Levine, M. S. Shadow enhancers as a source of evolutionary novelty. *Science* **321**, 1314 (2008).
111. Barolo, S. Shadow enhancers: frequently asked questions about distributed *cis*-regulatory information and enhancer redundancy. *Bioessays* **34**, 135–141 (2012).
112. Lin, Y. C. *et al.* Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nature Immunol.* **13**, 1196–1204 (2012).
113. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP–trait associations. *Nucleic Acids Res.* **42**, D1001–D1006 (2014).
114. Pasquali, L. *et al.* Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nature Genet.* **46**, 136–143 (2014).
115. Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S. & Snyder, M. Linking disease associations with regulatory information in the human genome. *Genome Res.* **22**, 1748–1759 (2012).
116. Trynka, G. *et al.* Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nature Genet.* **45**, 124–130 (2013).
117. Raj, T. *et al.* Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science* **344**, 519–523 (2014).
118. Kasowski, M. *et al.* Extensive variation in chromatin states across humans. *Science* **342**, 750–752 (2013).
119. Kilpinen, H. *et al.* Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. *Science* **342**, 744–747 (2013).
120. McVicker, G. *et al.* Identification of genetic variants that affect histone modifications in human cells. *Science* **342**, 747–749 (2013).
References 118–120 report on the effects of natural genetic variation in humans on the binding of transcription factors and histone modifications that are associated with enhancers and gene expression.
121. Degner, J. F. *et al.* DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* **482**, 390–394 (2012).
122. Gaffney, D. J. *et al.* Dissecting the regulatory architecture of gene expression QTLs. *Genome Biol.* **13**, R7 (2012).
123. Gaulton, K. J. *et al.* A map of open chromatin in human pancreatic islets. *Nature Genet.* **42**, 255–259 (2010).
124. Helgadottir, A. *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491–1493 (2007).
125. Cowper-Sal lari, R. *et al.* Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nature Genet.* **44**, 1191–1198 (2012).
126. Kasowski, M. *et al.* Variation in transcription factor binding among humans. *Science* **328**, 232–235 (2010).
127. Bauer, D. E. *et al.* An erythroid enhancer of *BCL11A* subject to genetic variation determines fetal hemoglobin level. *Science* **342**, 253–257 (2013).
128. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR–Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).

Acknowledgements
The authors were primarily supported by the US National Institutes of Health (NIH) grants DK091183, CA17390 and DK063491, and the San Diego Center for Systems Biology. C.E.R. was supported by the American Heart Association (12POST11760017) and the NIH Pathway to Independence Award (1K99HL12348).

Competing interests statement
The authors declare no competing interests.