Published in final edited form as:

Trends Genet. 2012 June ; 28(6): 276–284. doi:10.1016/j.tig.2012.02.008.

Enhancers as information integration hubs in development: lessons from genomics

Christa Buecker¹ and Joanna Wysocka^{1,2}

¹Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA, 94305, USA

²Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA, 94305, USA

Abstract

Transcriptional enhancers are the primary determinants of tissue-specific gene expression. Although the majority of our current knowledge of enhancer elements comes from detailed analyses of individual loci, recent progress in epigenomics has led to the development of methods for comprehensive and conservation-independent annotation of cell type-specific enhancers. Here, we discuss the advantages and limitations of different genomic approaches to enhancer mapping and summarize observations that have been afforded by the genome-wide views of enhancer landscapes, with a focus on development. We propose that enhancers serve as information integration hubs, at which instructions encoded by the genome are read in the context of a specific cellular state, signaling milieu and chromatin environment, allowing for exquisitely precise spatiotemporal control of gene expression during embryogenesis.

Keywords

enhancer; chromatin; epigenetic; development; growth factor signaling

Enhancers as mediators of crosstalk between the genome and its environment

During organismal development, cellular processes, such as proliferation, differentiation, adhesion and migration, allow for the emergence of complex forms and functions from a single set of instructions: the genome. Distinct cell states and behaviors are largely dictated by unique gene expression patterns, which arise as a consequence of the crosstalk between the genome and its environment. This crosstalk is mediated by the specialized *cis*-regulatory elements, including enhancers, silencers, promoters and insulators [1–4]. Among these, enhancers play a particularly prominent role in specifying spatiotemporal gene regulation during embryogenesis. The defining characteristic of enhancers is their ability to drive gene expression at a distance, independently of their orientation relative to the transcription start site [1,5,6] Modularity of enhancer elements, their uncoupling from promoters and

positioning away from transcription start sites, permits a given gene to be transcribed in multiple tissues, at varied levels, in response to distinct signaling cascades and at different stages during development via utilization of distinct, specialized enhancer modules [7,8].

Most of the current knowledge of tissue-specific *cis*-regulatory elements is based on detailed genetic and biochemical analyses of individual loci [8–10] (reviewed in [11]). Although these studies greatly improved the understanding of enhancer function, genome-wide identification of tissue-specific enhancers was precluded by the inability to annotate them systematically based on DNA sequence. Indeed, enhancers are compact, usually consisting of several hundred base pair-long stretches of DNA that are embedded within the vast noncoding genomic space and lack stereotypical sequence composition [5]. However, recent breakthroughs have come from advances in genomic technologies and the realization that genomic profiling of certain chromatin features associated with enhancers can be effectively used to map them in a genome-wide, cell type-specific and conservation-independent manner [12–19].

Here, we focus on synthesizing information from the genome-wide annotations of enhancer elements, with particular emphasis on analyses performed in embryonic cell types and in developmental contexts. Several excellent reviews recently covered in detail enhancer properties and existing models of enhancer-promoter communication [5,6,20]. For the purpose of our discourse, we give a brief overview of chromatin characteristics associated with enhancers. We then summarize how these different properties can be harnessed for annotation of cell type-specific enhancer elements, and we compare different strategies for epigenomic enhancer mapping. Finally, we discuss insights into enhancer abundance and function that have been revealed through genomic studies. Resulting global views not only generalized many observations previously made at individual enhancers, but also revealed interesting surprises, such as extreme dynamics in enhancer utilization patterns in different cell types and widespread priming of developmental enhancers [14,16,18,19].

Properties of enhancer chromatin

Common properties of enhancer chromatin are visualized in Figure 1a. A central feature of enhancers is their ability to function as integrated transcription factor (TF)-binding platforms. Consequently, enhancers commonly contain clustered recognition sites for multiple TFs representing distinct classes of DNA binding proteins. Indeed, genome-wide studies of TF occupancy reveal that most lineage-specific TF-binding events occur not at proximal promoters, but at distal regulatory regions located at distances ranging from one to hundreds or, in rare instances, even thousands of kilobases away from the nearest transcription start site (TSS). Moreover, in a given cell type, core members of the regulatory network show extensive genomic co-occupancy and often cooperate in their binding to DNA. For example, the core TFs in embryonic stem cells (ESCs), Oct4, Sox2 and Nanog, share a large portion of genomic targets, with >90% of co-binding events occurring outside of proximal promoters at regions exhibiting enhancer features [21–23]. The presence of TFs at enhancers is associated with regions of low nucleosome occupancy, which exhibit high sensitivity to DNA nucleases (Figure 1a) [24–26]. Rather than being entirely nucleosome

free, DNase hypersensitive regions appear to dynamically bind highly mobile nucleosomes containing the specialized histone variants H3.3 and H2A.Z [27].

The ability of TFs to activate transcription is dependent on the recruitment of coactivator proteins, which often lack sequence-specific DNA-binding competency and function as histone modifiers (e.g. acetyltransferases p300/CBP, Gcn5-containing ATAC complex), ATP-dependent chromatin remodelers catalyzing nucleosome movement (e.g. chromodomain helicase DNA binding protein 7 (CHD7), Brg1 complex BAF), or mediators of crosstalk with basal transcriptional machinery at promoters (e.g. Mediator complex) [28–31]. Whereas combinations of TFs bound at enhancers are cell type specific, most individual TFs, irrespective of cell type, can recruit coactivators, which are often broadly expressed. As such, at least one TF within the cluster is likely to recruit a given coactivator; therefore, genome-wide profiling of coactivator occupancy provides a powerful strategy for enhancer detection that does not require any prior knowledge of the core circuitry governing expression in a particular lineage (note overlap of TF and p300 enrichments at representative loci in Figure 1b). Consistent with this, many of the aforementioned coactivators were found to be broadly associated with enhancer elements in genomic studies, although it is likely that no single coactivator can capture a complete enhancer repertoire [14,16,32–36] (Figure 1).

A first indication that distal regulatory elements may be commonly marked by specific histone modifications came from analyses of the 1% of the human genome via the ENCODE project, which linked the presence of histone H3 lysine 4 monomethylation (H3K4me1) to distal enhancer regions [12,37]; these observations were subsequently confirmed by genome-wide studies in different species and varied cell types [14,34,38,39]. Interestingly, developmental enhancers can exist in active, as well as poised states [Figure 1a, compare (i) and (ii)], which can be distinguished by the presence of histone H3 lysine 27 acetylation (H3K27ac) at active enhancers [16,19], discussed in more detail below. Active enhancers flanked by acetylated nucleosomes are able to recruit RNA polymerase II (Pol II), leading to the production of enhancer-originating, short bidirectional RNAs of unknown function, termed eRNAs [40–42]. Other histone modifications have been linked to active [e.g. histone H3 acetylated Lys9 (H3K9ac) and histone H3 trimethyl Lys79 (H3K79me3) [43,44]] and poised [e.g. histone H3 trimethyl Lys27 (H3K27me3) and histone H3 trimethyl Lys9 (H3K9me3) [16,18]] enhancer identities in genomic studies. However, it is important to note that none of the aforementioned modifications alone are unique to enhancer regions. For example, H3K4me1 is abundant downstream of TSSs within gene bodies, whereas H3K27ac and H3K9ac are also enriched at proximal promoter regions. Furthermore, given the low nucleosomal density at enhancers, nucleosomes marked with H3K4me1 and H3K27ac flank, rather than directly overlap, TF and coactivator-occupied enhancer regions (e.g. compare genomic tracts for p300 and TFs with those of histone modifications in Figure 1b).

Strategies for genomic annotation of enhancer elements

In principle, any of the properties of enhancer chromatin (Figure 1), including overlap with low nucleosomal density sites, co-occupancy of multiple TFs, binding of coactivators, enrichment for specific combinations of histone modifications and Pol II association with production of eRNAs, can provide global insights into enhancer genomic positions when

assayed genome-wide in a specific cell type. Different strategies have been recently used in conjunction with high-throughput sequencing methods to uncover regulatory regions in the genome, demonstrating advantages and limitations of each approach (Table 1). Given the combinatorial nature of gene expression regulation and the complexity of animal development, we favor the view that no single modification or factor, but rather integration of multiple layers of genomic information, will yield the most comprehensive annotation and understanding of developmental enhancers. In support of this view, combining patterns of multiple histone modifications with coactivator or Pol II occupancy predicts enhancer position and activity with high accuracy and resolution [16,44]. Moreover, rapid progress in genomics is already leading to the development of alternative methods of enhancer discovery, some of which do not rely on chromatin profiling. For example, a recent study suggests that detection of short bidirectional eRNA transcripts can be used for the prediction of active enhancer elements genome wide [42].

One challenging issue that remains to be addressed is genome-wide mapping of enhancer–promoter interactions. Current methodologies of evaluating physical proximity between distant chromatin regions are based on the derivatives of the Chromosome Conformation Capture assay, such as 3C, 4C and 5C (reviewed in [45,46]). These are low-to medium-throughput assays appropriate for assessing interactions between two select loci (3C), one locus and the rest of the genome (4C) or all interactions between multiple preselected elements (5C) [45,46]. Despite limited throughput, these assays have been successfully used both for the identification of novel regulatory regions, as well as for the high-resolution mapping of interactions within a chromosomal domain ([47], Gheldorf). Moreover, a newly developed variant of the Chromosome Conformation Capture, termed Hi-C, allows for genome-wide, unbiased mapping of long-range interactions [48]. Although the current resolution of the Hi-C technology (approximately 0.1–1 Mb for mammalian genomes) makes it unsuitable for precise analysis of enhancer–promoter interactions, this resolution is limited mostly by the sequencing depth and its associated costs. Chromosomal interactions within smaller metazoan genomes, such as that of *Drosophila*, are already being analyzed by Hi-C at much higher resolution [49], and we expect that such analyses will soon become possible for mammalian genomes. Also lagging behind the staggering amount of genomic data is the ability to examine spatio-temporal activity of predicted enhancers in a high-throughput, unbiased and dynamic way in the context of the developing embryo, although resources aimed at relatively broad validation of regulatory regions in mouse embryos, such as the Vista Enhancer Browser (<http://enhancer.lbl.gov>) are being developed.

Enhancers are the primary determinant of tissue-specific gene expression

Regardless of specific annotation methods, genomic analyses generally support the long-held view that enhancers are the primary determinant of tissue-specific gene expression. Consistently, p300 occupancy and enhancer-associated histone modification patterns are highly cell-type specific compared with promoter-marking patterns or occupancy of the insulator-associated protein CCCTC-binding factor (CTCF) [12,14,43]. What has been truly eye opening is the number of sites marked by active enhancer signatures in each cell type (ranging from thousands to tens of thousands, depending on the annotation method and reagents), and the extent of the developmental dynamics in enhancer utilization patterns

[12,14,43]. Even when related, genetically matched embryonic cell types are compared, such as undifferentiated human ESCs (hESCs) and those induced to form early mesendoderm or neuroepithelium, most enhancer chromatin signatures are specific to each cell type, whereas promoter-associated modifications remain largely invariant [50] (A. Rada-Iglesias and J. Wysocka, unpublished).

These observations have interesting implications for the view of enhancer element contribution to the genome. During the developmental process, each of the over 200 specialized cell types present in the human body originates from pluripotent cells of the early embryo through differentiation progressing via many, often transient, stem and progenitor cell states. Although it is difficult to define and measure how many such transient cell states arise during development for all cell types combined, the figure is likely to be at least in the thousands. Thus, given that between hESCs, early mesendoderm and early neuroepithelium alone there are thousands of enhancer signatures specific for each cell state, the number of unique regulatory elements harbored by the human genome may well exceed a million, even if a substantial portion of the early enhancers is reutilized at later developmental stages or in adulthood. Of course, such estimates are never precise, although a similar figure has been suggested by others [12]. Nevertheless, if we assume that this estimate is close to reality, and consider a typical enhancer size of 200–500 bp, as much as 10% of the human genome may encode enhancer elements.

Why did metazoan genomes acquire such an astounding number of regulatory elements? Changes in *cis*-regulatory sequences are a major mechanism underlying morphological evolution [51,52]. Expansion and functional diversification of regulatory elements allows for enormous combinatorial complexity of expression patterns and contributes to the emergence of intricate cell behaviors without a large expansion of the protein-coding gene repertoire in metazoan species [1,7,53,54]. Indeed, many key developmental genes are ‘recycled’ numerous times during embryogenesis through cell type-specific enhancer-mediated regulation. For example, *Hox* genes play a fundamental role in the anterior–posterior patterning of the animal body across metazoans. However, during the evolution of tetrapods, *Hox* genes have been coopted to orchestrate limb patterning and digit formation through evolution of novel enhancer elements modulating their expression in the limb [55,56,57]. In addition to enhancers driving transcription of protein-coding genes in different tissues, it is becoming clear that a large number of regulatory elements controlling cell type-specific expression of noncoding RNA genes must exist [58,59], further expanding enhancer repertoire.

Moreover, even within a given cell type, each active developmental locus is commonly associated with multiple enhancers, suggesting that information from several enhancers can be combinatorially read by a single promoter. Although this view still needs to be substantiated on a genome-wide level, a recent elegant study demonstrated that *Hox* gene expression in digits is regulated by multiple enhancers positioned within a ‘regulatory archipelago’, with each element contributing either quantitatively or qualitatively to transcription of the *HoxD* cluster genes [60]. In addition to combinatorial integration of multiple inputs, some enhancers appear to have redundant functions. In *Drosophila*, so-called ‘shadow enhancers’ show cell-type specificity and TF-binding properties consistent

with that of the ‘primary’ enhancer, but are located more remotely from the TSS of the target gene [61]. Shadow enhancers were suggested to confer redundancy and phenotypic robustness, although it remains unclear how many of these elements are truly redundant as opposed to essential under conditions not yet tested [61,62].

Widespread phenomenon of enhancer priming

One of the surprises following the genomic mapping of regulatory elements was the widespread pre-patterning of developmental enhancers. In human and mouse ESCs, the enhancers of genes associated with early differentiation are pre-marked by a ‘poised’ chromatin signature, which shares many properties of active enhancers, such as low nucleosomal density, the presence of TFs and coactivators, and enrichment of H3K4me1 (Figure 1) [16,18,19]. However, in contrast to active enhancers, the H3K27 residue at nucleosomes flanking poised enhancers is not acetylated, instead it is often modified by Polycomb-mediated methylation [16,18,19] (Figure 1). Widespread epigenetic priming of developmental enhancers in ESCs is probably an important mechanism underlying anticipation of future developmental fates and, thus, pluripotency of these cells. Notably, enhancer priming is not limited to ESCs [19], although enhancer-associated H3K27me3 appears most prevalent in undifferentiated cells. Moreover, some H3K4me1-marked regions represent enhancers that were active at preceding developmental stages and removal of histone methylation significantly lags behind the disappearance of primary TFs and coactivators from enhancers [48].

In the poised state, enhancers lack the ability to drive gene expression [16]. Consistently, genes in proximity to poised enhancers are inactive and commonly have so-called ‘bivalently marked’ promoters, characterized by the simultaneous presence of H3K4me3 and H3K27me3 at the TSS [16,18,63]. These observations suggest that enhancer and promoter priming are coordinated. One example of the mechanism underlying such coordination comes from a recent study of the murine myogenic differentiation (*MyoD*) locus during reprogramming that demonstrated that binding of a reprogramming TF, Oct4, to its DNA recognition site within the *MyoD* enhancer leads to the establishment of a poised signature, which in turn results in the poised, bivalent promoter marking pattern [64].

During differentiation into a specific lineage, a cell type-specific subset of poised enhancers acquires H3K27ac, Pol II association and eRNAs [16,44]. The switch to this active enhancer signature coincides with the gain of ability to drive gene expression. Developmental decisions are mediated by signaling pathways; therefore, transitions between poised and active enhancer states must be closely integrated with the signaling environment. Several recent reports support this notion. For example, the active state of the well-studied *Nodal* enhancer requires H3K27me3 demethylase JMJD3, whose recruitment to the enhancer is dependent on Smad2/3, transcriptional effectors of Nodal–Activin signaling [65]. Similarly, in breast cancer cells, the estrogen receptor recruits JMJD3 to the B-cell CLL/lymphoma 2 (*BCL2*) enhancer in a hormone-dependent manner, leading to its transition from a poised to an active state [66]. Another study demonstrated the requirement for the enhancer-mediated recruitment of JMJD3 in the eviction of the repressive Polycomb repressive complex 2 (PRC2) from the poised α -globin promoter in erythroid cells [67]. Finally, during heart

development, UTX, the second known H3K27me3 demethylase, is recruited to poised heart-specific enhancers and is necessary for the switch to an active enhancer state [68]. If these observations from individual enhancer–promoter pairs can be generalized, signaling-dependent recruitment of H3K27me3 demethylases to enhancer regions may prove to be a common mechanism by which Polycomb-mediated repression is coordinately counteracted at poised enhancers and their target promoters.

Interestingly, p300 acetyltransferase is present both at active and poised enhancers, but only active enhancers are acetylated at H3K27, a major substrate residue for p300/BP [69,70]. These observations suggest that mechanisms counteracting acetylation, either through direct inhibition of p300 enzymatic activity or facilitating turnover through recruitment of histone deacetylases (HDACs) or via a mutually exclusive relationship with H3K27me3 and H3K9me3, must operate at poised enhancers to prevent their premature activation.

Role of TFs in enhancer priming

Widespread occupancy of TFs occurs not only at active, but also at poised enhancers [19,71]. For example, over half of all poised (and a similar proportion of active) enhancers are bound by OCT4 in hESCs. These observations have interesting implications for the role played by the core TFs of a given cell type in setting the stage for future developmental events. In case of ESCs, it has been suggested that core TFs, often referred to as pluripotency factors, play a role in directing differentiation [72,73]. What are the mechanisms by which TFs participate in enhancer priming? Existing literature points towards several potential mechanisms, including what we refer to here as the ‘placeholder model’ and the ‘tug-of-war model’ (Figure 2a). Generally, the placeholder model relies on the fact that most TFs belong to large families, such as Sox, POU, or Forkhead, characterized by homologous DNA-binding domains sharing similar DNA recognition motifs. Different members of a given TF family are commonly expressed in a cell type-specific manner, and may have distinct, context-dependent trans-activation properties. During a cell fate transition, one factor (expressed in either an ESC or progenitor cell) is replaced by another family member (expressed in the differentiating cell) at the same DNA recognition motif residing within an enhancer (Figure 2a). In support of this model, in ESCs, Sox2 associates with many poised enhancers destined to be bound and activated by Sox3 in neural progenitors [74]. Similarly, Sox2 primes a pre-B cell specific enhancer, which is ultimately activated by Sox4 in pre-B cells [75], whereas forkhead box D3 (FoxD3) associates with *Alb1* endodermal enhancer in ESCs, but is replaced by FoxA1 following differentiation to endoderm [76].

By contrast, the ‘tug-of-war’ model stipulates that pluripotency factors are lineage specifiers that promote activation of specific fates while repressing alternative fates (Figure 2b) [73]. In this context, Oct4 was proposed to have a pivotal role in the establishment of the mesendoderm layer, whereas Sox2 is crucial for the induction of the neuroectodermal lineage. The undifferentiated state is maintained via broad co-occupancy of pluripotency factors at developmental enhancers, resulting in the ‘tug-of-war’ and preventing premature activation [73]. One example of such a mechanism is the regulation of the *eomesodermin* (*EOMES*) enhancer, which is bound by OCT4, SOX2 and NANOG in hESCs but remains in

a poised state [77]. At the onset of endoderm specification, activation of this enhancer is dependent on the departure of SOX2 and the presence of NANOG and is subsequently augmented by the protein product of the enhancer target gene, *EOMES* [77].

Taken together, multiple lines of evidence illuminate the important role of TFs in enhancer priming. The placeholder and tug-of-war models discussed here are not mutually exclusive, and both require further studies. For example, it remains unclear what determines context-dependent activity of TFs at poised enhancers, and how this relates to the chromatin features associated with these elements.

It is also important to note that not all developmental enhancers are pre-marked in pluripotent cells. Instead, many enhancers are created *de novo* during differentiation, in a process that, at least in some instances, is dependent on the so-called ‘pioneer factors’, which are often lineage specific and belong to the core transcriptional circuitry [78]. However, in contrast to most TFs, pioneer factors have the ability to bind nucleosome-wrapped DNA [78]. These pioneering associations lead to subsequent recruitment of chromatin remodelers and establishment of low nucleosomal density regions, creating a window of opportunity for other TFs and signaling effectors to bind DNA, in turn allowing for the developmental dynamics in enhancer utilization patterns [78].

Enhancers as information integration hubs

In a simplified view, most signaling cascades rely on transducing information from a ligand-bound membrane receptor to a DNA-binding effector, which upon activation and translocation to the nucleus confers specific transcriptional outputs. Genome-wide studies support earlier observations that, in addition to serving as binding platforms for core TFs, enhancers are occupied by DNA-binding effectors of cellular signaling pathways [22,79,80]. Most signaling effectors lack pioneer factor ability and so their binding is restricted to the existing low nucleosomal density regions, delineated by the core circuitry. Recent reports support an even closer relationship between core lineage regulators and signaling pathway output: the genomic binding of transcriptional effectors of TGF β , BMP and Wnt pathways not only follows that of a master regulator of the lineage identity, but can be broadly redirected by the overexpression of a core factor for another cell type [79,80]. These observations can in part explain why core factors have the ability to establish the cellular identity even in a cell of different origin and state, leading to reprogramming and transdifferentiation [81–85].

Thus, enhancers serve as information integration hubs, where genomic instructions (dictated by the underlying sequence), cellular environment (defined by the core circuitry), signaling environment (defined by the signaling effectors) and epigenetic information (dictated by the restrictive or permissive chromatin states) are brought together to establish cell type-specific gene expression. Clearly, another layer of complexity, which has been discussed briefly above, is associated with enhancer–promoter communication. It is easy to envision that this additional level of combinatorial regulation may allow input from multiple enhancers to be integrated at a single promoter [60]. Conversely, perhaps information from a single enhancer may be conveyed at multiple promoters for coordinated gene regulation. Extensive

interactions between transcriptionally active promoters have been reported in human cells [86], potentially corresponding to the so-called ‘transcription factories’, the nuclear foci containing active RNA polymerase [87]. If gene expression is indeed organized in transcription factories, it remains to be discovered how enhancers regulate the inclusion and/or exclusion of specific genes or gene clusters from such factories to allow for precise control of gene expression.

Concluding remarks

Recent advances in high-throughput sequencing technologies enabled the replacement of painstaking mapping of single enhancers at individual loci with genome-wide epigenomic annotations based on various common features of enhancer chromatin. As a result, the catalogue of cell type-specific regulatory regions has expanded immensely in recent years, although only a miniscule fraction of these elements has been validated and characterized *in vivo*. Current data permit generalization of many observations made in earlier studies of individual loci and show that enhancers are the most dynamically utilized part of the human genome, with the ‘junk DNA’ of one cell type being the treasure of another cell type. Moreover, it is becoming clear that enhancers serve as signal integration hubs, at which genetic information is read in the context of unique cellular and signaling environments that are specified by lineage TFs, signaling effectors and chromatin modifiers. Nevertheless, many outstanding questions remain; for example, what are the mechanisms by which enhancers find and communicate with their target promoters in the vast genomic space?

Importantly, the wealth of current epigenomic data provides information that would be difficult to deduce from analyses of single loci. For example, knowledge of thousands of elements uniquely marked by the enhancer signatures in a specific cell type affords the statistical power to identify enriched sequence motifs and predict new core TFs in a given circuitry. First forays demonstrating a strong predictive power of such an approach have already been made [88]. Finally, although it is well documented that enhancer mutations are linked to various human pathologies [89,90], their role has not been systematically examined in the past, largely because of the scarcity of information on relevant genomic sequences. However, epigenomic annotation of tissue-specific enhancers can readily define regulatory space that can be subsequently searched for disease-associated sequence variants identified through genome-wide association studies. Initial analyses of this type show that disease-linked single nucleotide polymorphisms are indeed frequently positioned within cell-type specific enhancer elements [43,91], suggesting a large universe of noncoding regulatory mutations in human disease that can be explored at last.

Acknowledgments

We thank A. Rada-Iglesias, R. Greenberg and anonymous reviewers for comments on the manuscript. Work in our laboratory is funded by CIRM (RB3-05100, RN1 00579-1), NIH (GM095555) and Keck Foundation.

References

1. Bulger M, Groudine M. Enhancers: the abundance and function of regulatory sequences beyond promoters. *Dev Biol.* 2010; 339:250–257. [PubMed: 20025863]

2. Bushey AM, et al. chromatin insulators: regulatory mechanisms and epigenetic inheritance. *Mol Cell*. 2008; 32:1–9. [PubMed: 18851828]
3. Riethoven JJM. Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Methods Mol Biol*. 2010; 674:33–42. [PubMed: 20827584]
4. Noonan JP, McCallion AS. Genomics of long-range regulatory elements. *Annu Rev Genomics Hum Genet*. 2010; 11:1–23. [PubMed: 20438361]
5. Ong CT, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet*. 2011; 12:283–293. [PubMed: 21358745]
6. Williamson I, et al. Enhancers: from developmental genetics to the genetics of common human disease. *Dev Cell*. 2011; 21:17–19. [PubMed: 21763601]
7. Uchikawa M, et al. Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev Cell*. 2003; 4:509–519. [PubMed: 12689590]
8. Granier C, et al. Nodal *cis*-regulatory elements reveal epiblast and primitive endoderm heterogeneity in the peri-implantation mouse embryo. *Dev Biol*. 2011; 349:350–362. [PubMed: 21047506]
9. Yeom YI, et al. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development*. 1996; 122:881–894. [PubMed: 8631266]
10. Norris DP, Robertson EJ. Asymmetric and node-specific nodal expression patterns are controlled by two distinct *cis*-acting regulatory elements. *Genes Dev*. 1999; 13:1575–1588. [PubMed: 10385626]
11. Maston GA, et al. Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet*. 2006; 7:29–59. [PubMed: 16719718]
12. Heintzman ND, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007; 39:311–318. [PubMed: 17277777]
13. Heintzman ND, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009; 459:108–112. [PubMed: 19295514]
14. Visel A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*. 2009; 457:854–858. [PubMed: 19212405]
15. May D, et al. Large-scale discovery of enhancers from human heart tissue. *Nat Genet*. 2011; 44:89–93. [PubMed: 22138689]
16. Rada-Iglesias A, et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2010; 470:279–283. [PubMed: 21160473]
17. Pennacchio LA, Visel A. Limits of sequence and functional conservation. *Nat Genet*. 2010; 42:557–558. [PubMed: 20581875]
1. Zentner GE, et al. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res*. 2011; 21:1273–1283. [PubMed: 21632746]
19. Creighton MP, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci USA*. 2010; 107:21931–21936. [PubMed: 21106759]
20. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. *Cell*. 2011; 144:327–339. [PubMed: 21295696]
21. Young RA. Control of the embryonic stem cell state. *Cell*. 2011; 144:940–954. [PubMed: 21414485]
22. Chen X, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*. 2008; 133:1106–1117. [PubMed: 18555785]
23. Boyer LA, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*. 2005; 122:947–956. [PubMed: 16153702]
24. He HH, et al. Nucleosome dynamics define transcriptional enhancers. *Nat Genet*. 2010; 42:343–347. [PubMed: 20208536]
25. Mito Y, et al. Histone replacement marks the boundaries of *cis*-regulatory domains. *Science*. 2007; 315:1408–1411. [PubMed: 17347439]
26. Song L, et al. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res*. 2011; 21:1757–1767. [PubMed: 21750106]

27. Jin C, et al. H3.3/H2A.Z. double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet.* 2009; 41:941–945. [PubMed: 19633671]
28. Roeder RG. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett.* 2005; 579:909–915. [PubMed: 15680973]
29. D'Alessio JA, et al. Shifting players and paradigms in cell-specific transcription. *Mol Cell.* 2009; 36:924–931. [PubMed: 20064459]
30. Weake VM, Workman JL. Inducible gene expression: diverse regulatory mechanisms. *Nat Rev Genet.* 2010; 11:426–437. [PubMed: 20421872]
31. Borggrefe T, Yue X. Interactions between subunits of the Mediator complex with gene-specific transcription factors. *Semin Cell Dev Biol.* 2011; 22:759–768. [PubMed: 21839847]
32. Schnetz MP, et al. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet.* 2010; 6:e1001023. [PubMed: 20657823]
33. Kagey MH, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature.* 2010; 467:430–435. [PubMed: 20720539]
34. Ghisletti S, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity.* 2010; 32:317–328. [PubMed: 20206554]
35. Krebs AR, et al. SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Mol Cell.* 2011; 44:410–423. [PubMed: 22055187]
36. Ram O, et al. Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell.* 2011; 147:1628–1639. [PubMed: 22196736]
37. Birney E, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature.* 2007; 447:799–816. [PubMed: 17571346]
38. Nègre N, et al. A *cis*-regulatory map of the *Drosophila* genome. *Nature.* 2011; 471:527–531. [PubMed: 21430782]
39. Aday AW, et al. Identification of *cis* regulatory features in the embryonic zebrafish genome through large-scale profiling of H3K4me1 and H3K4me3 binding sites. *Dev Biol.* 2011; 357:450–462. [PubMed: 21435340]
40. Kim TK, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature.* 2010; 465:182–187. [PubMed: 20393465]
41. Wang D, et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature.* 2011; 474:390–394. [PubMed: 21572438]
42. Melgar MF, et al. Discovery of active enhancers through bidirectional expression of short transcripts. *Genome Biol.* 2011; 12:R113. [PubMed: 22082242]
43. Ernst J, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature.* 2011; 473:43–49. [PubMed: 21441907]
44. Bonn S, et al. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet.* 2012; 44:148–156. [PubMed: 22231485]
45. van Steensel B, Dekker J. Genomics tools for unraveling chromosome architecture. *Nat Biotechnol.* 2010; 28:1089–1095. [PubMed: 20944601]
46. de Wit E, de Laat W. A decade of 3C technologies: insights into nuclear organization. *Genes Dev.* 2012; 26:11–24. [PubMed: 22215806]
47. Gheldof N, et al. Cell-type-specific long-range looping interactions identify distant regulatory elements of the *CFTR* gene. *Nucleic Acids Res.* 2010; 38:4325–4336. [PubMed: 20360044]
48. Lieberman-Aiden E, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009; 326:289–293. [PubMed: 19815776]
49. Sexton T, et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell.* 2012; 148:458–472. [PubMed: 22265598]
50. Hawkins RD, et al. Dynamic chromatin states in human ES cells reveal potential regulatory sequences and genes involved in pluripotency. *Cell Res.* 2011; 21:1393–1409. [PubMed: 21876557]

51. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*. 2008; 134:25–36. [PubMed: 18614008]
52. Frankel N, et al. Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature*. 2011; 474:598–603. [PubMed: 21720363]
53. Levine M. Transcriptional enhancers in animal development and evolution. *Curr Biol*. 2010; 20:R754–R763. [PubMed: 20833320]
54. Arnold K, et al. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. *Cell Stem Cell*. 2011; 9:317–329. [PubMed: 21982232]
55. Spitz F, et al. A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell*. 2003; 113:405–417. [PubMed: 12732147]
56. Schneider I, et al. Appendage expression driven by the Hoxd Global Control Region is an ancient gnathostome feature. *Proc Natl Acad Sci USA*. 2011; 108:12782–12786. [PubMed: 21765002]
57. Woltering JM, Duboule D. The origin of digits: expression patterns versus regulatory mechanisms. *Dev Cell*. 2010; 18:526–532. [PubMed: 20412768]
58. Pauli A, et al. Non-coding RNAs as regulators of embryogenesis. *Nat Rev Genet*. 2011; 12:136–149. [PubMed: 21245830]
59. Guttman M, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*. 2011; 477:295–300. [PubMed: 21874018]
60. Montavon T, et al. A regulatory archipelago controls *Hox* genes transcription in digits. *Cell*. 2011; 147:1132–1145. [PubMed: 22118467]
61. Barolo S. Shadow enhancers: frequently asked questions about distributed *cis*-regulatory information and enhancer redundancy. *Bioessays*. 2012; 34:135–141. [PubMed: 22083793]
62. Frankel N, et al. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature*. 2010; 466:490–493. [PubMed: 20512118]
63. Bernstein BE, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2006; 125:315–326. [PubMed: 16630819]
64. Taberlay PC, et al. Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell*. 2011; 147:1283–1294. [PubMed: 22153073]
65. Dahle Ø, et al. Nodal signaling recruits the histone demethylase Jmjd3 to counteract polycomb-mediated repression at target genes. *Sci Signal*. 2010; 3:ra48. [PubMed: 20571128]
66. Svtelis A, et al. H3K27 demethylation by JMJD3 at a poised enhancer of anti-apoptotic gene *BCL2* determines ERα ligand dependency. *EMBO J*. 2011; 30:3947–3961. [PubMed: 21841772]
67. Vernimmen D, et al. Polycomb eviction as a new distant enhancer function. *Genes Dev*. 2011; 25:1583–1588. [PubMed: 21828268]
68. Lee S, et al. UTX, a histone h3-lysine 27 demethylase, acts as a critical switch to activate the cardiac developmental program. *Dev Cell*. 2011; 22:25–37. [PubMed: 22192413]
69. Jin Q, et al. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J*. 2011; 30:249–262. [PubMed: 21131905]
70. Pasini D, et al. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res*. 2010; 38:4958–4969. [PubMed: 20385584]
71. Heinz S, et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010; 38:576–589. [PubMed: 20513432]
72. Loh KM, Lim B. A precarious balance: pluripotency factors as lineage specifiers. *Cell Stem Cell*. 2011; 8:363–369. [PubMed: 21474100]
73. Thomson M, et al. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell*. 2011; 145:875–889. [PubMed: 21663792]
74. Bergsland M, et al. Sequentially acting Sox transcription factors in neural lineage development. *Genes Dev*. 2011; 25:2453–2464. [PubMed: 22085726]
75. Liber D, et al. Epigenetic priming of a pre-B cell-specific enhancer through binding of Sox2 and Foxd3 at the ESC stage. *Cell Stem Cell*. 2010; 7:114–126. [PubMed: 20621055]

76. Xu J, et al. Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. *Genes Dev.* 2009; 23:2824–2838. [PubMed: 20008934]
77. Teo AKK, et al. Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev.* 2011; 25:238–250. [PubMed: 21245162]
78. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 2011; 25:2227–2241. [PubMed: 22056668]
79. Trompouki E, et al. Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. *Cell.* 2011; 147:577–589. [PubMed: 22036566]
80. Mullen AC, et al. Master transcription factors determine cell-type-specific responses to TGF- β signaling. *Cell.* 2011; 147:565–576. [PubMed: 22036565]
81. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* 2006; 126:663–676. [PubMed: 16904174]
82. Vierbuchen T, et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature.* 2010; 463:1035–1041. [PubMed: 20107439]
83. Hochedlinger K, Plath K. Epigenetic reprogramming and induced pluripotency. *Development.* 2009; 136:509–523. [PubMed: 19168672]
84. Ieda M, et al. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell.* 2010; 142:375–386. [PubMed: 20691899]
85. Szabo E, et al. Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature.* 2010; 468:521–526. [PubMed: 21057492]
86. Li G, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell.* 2012; 148:84–98. [PubMed: 22265404]
87. Cook PR. A model for all genomes: the role of transcription factories. *J Mol Biol.* 2010; 395:1–10. [PubMed: 19852969]
88. Verzi MP, et al. Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Dev Cell.* 2010; 19:713–726. [PubMed: 21074721]
89. Kleinjan DJ, Coutinho P. *Cis*-ruption mechanisms: disruption of *cis*-regulatory control as a cause of human genetic disease. *Brief Funct Genomic Proteomic.* 2009; 8:317–332. [PubMed: 19596743]
90. VanderMeer JE, Ahituv N. *Cis*-regulatory mutations are a genetic cause of human limb malformations. *Dev Dyn.* 2011; 240:920–930. [PubMed: 21509892]
91. Teng L, et al. Enhancers in embryonic stem cells are enriched for transposable elements and genetic variations associated with cancers. *Nucleic Acids Res.* 2011; 39:7371–7379. [PubMed: 21685454]
92. Kunarso G, et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat Genet.* 2010; 42:631–634. [PubMed: 20526341]
93. Hanna J, et al. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad Sci USA.* 2010; 107:9222–9227. [PubMed: 20442331]
94. Tesar PJ, et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature.* 2007; 448:196–199. [PubMed: 17597760]
95. Giresi PG, et al. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.* 2007; 17:877–885. [PubMed: 17179217]
96. Nagy PL, Price DH. Formaldehyde-assisted isolation of regulatory elements. *Wiley Interdiscip Rev Syst Biol Med.* 2009; 1:400–406. [PubMed: 20046543]
97. Wilhelm BT, et al. Differential patterns of intronic and exonic DNA regions with respect to RNA polymerase II occupancy, nucleosome density and H3K36me3 marking in fission yeast. *Genome Biol.* 2011; 12:R82. [PubMed: 21859475]
98. Berchowitz LE, et al. A positive but complex association between meiotic double-strand break hotspots and open chromatin in *Saccharomyces cerevisiae*. *Genome Res.* 2009; 19:2245–2257. [PubMed: 19801530]

99. Zinzen RP, et al. Combinatorial binding predicts spatio-temporal *cis*-regulatory activity. *Nature*. 2009; 462:65–70. [PubMed: 19890324]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Chromatin properties at active and poised enhancers. **(a)** Schematic representation of proteins, histone modifications and RNA found at active (i) and poised (ii) enhancers. An active, but not poised, enhancer has the ability to drive gene expression. At both enhancer classes, multiple transcription factors (TF1 and TF2, orange), DNA-binding active signaling effectors (aSE, blue) and coactivators (p300, purple) occupy the central region of low nucleosomal density, which is hypersensitive to DNase. In addition, active enhancers are bound by RNA-polymerase II (Pol II, light green) which produces bidirectional short RNAs called eRNAs. By contrast, poised enhancers lack Pol II, but, at least in human embryonic stem cells (hESCs), are occupied by the Polycomb repressive complex 2 (PRC2, yellow). The nucleosomes flanking enhancer regions are marked by monomethylation of histone H3 lysine 4 (H3K4me1, light blue). Lysine 27 of histone H3 is commonly acetylated at the nucleosomes flanking active enhancers (H3K27ac, dark green) but methylated at poised enhancers (H3K27me3, red). **(b)** Genome browser representations of select protein and histone modification enrichments at a model loci containing active (*POU5F1/OCT4*, left) and poised (*EOMES*, right) enhancers (box) in hESCs. WIG files from published data [16,80,92] for p300 (coactivator, purple), SMAD3 (active signaling effector, blue), OCT4 (TF, orange), NANOG (TF, orange), H3K4me1 (light blue), H3K27ac (green) and H3K27me3 (red) were generated using QuEST and imported into the UCSC browser. Note the tight overlap of TF (OCT4, NANOG and SMAD3) and p300 binding, and broader regions surrounding the enhancers and showing H3K4me1 enrichments. H3K27ac (green) flanks active enhancers, but is completely absent at the poised enhancer where the same lysine residue is methylated over a broader chromosomal region (red). OCT4 expression is driven by two conserved enhancers, the distal (DE) and the proximal enhancer (PE), with a distinct activity during early embryonic development; both enhancers are active in human embryonic stem cells [9,93,94].

Figure 2.

Models of transcription factor (TF) involvement in enhancer priming. In the 'placeholder' model (a), the poised enhancer is established during development by TF1 (dark green) binding to its recognition motive (RM, orange); other TFs or active signaling effectors can also be recruited in the poised state. During differentiation, TF1 is replaced by a different member of the same TF family (here TF2), which also recognizes RM. This replacement leads to the activation of the enhancer [acetylation of H3K27 (K27ac, green)] and target gene activation. (b) The 'tug-of-war' model proposes that, during the establishment and maintenance of the poised enhancers, TFs with opposing activities (TF3, pink and TF4, orange) are bound to the same enhancer. TF3 is able to activate the enhancer and with it the expression of the associated gene, whereas TF4 is suppressing this activation. During differentiation the expression (or chromatin association) of the opposing TF4 is diminished, allowing TF3 to activate the enhancer and transcription of the gene target. These models are not mutually exclusive; indeed, most instances of gene activation are probably tightly regulated by a combination of these models together with additional signaling effectors and TFs that become induced during the differentiation process.

Table 1
Summary of epigenomic methods to map enhancers

Approach	Method	Advantages	Limitations	Refs
Low nucleosomal density-based mapping	Formaldehyde-assisted isolation of regulatory elements (FAIRE) or DNase hypersensitivity coupled with high-throughput sequencing (DNase-seq)	Technically simple	Noisy, detects a large variety of regulatory regions associated with depleted or highly mobile nucleosomes: enhancers, insulators, promoters and terminators, 3' ends of introns, etc.	[26,95–98]
		Can be performed on small numbers of cells (or even single cells), as no enrichment through protein-based method is required		
Clustered TF occupancy-based mapping	Chromatin immunoprecipitation (ChIP)-seq of multiple core TFs for a given lineage, followed by identification of clustered TF-binding sites	Clustered co-occupancy of TFs is a key feature of enhancers; thus, this method predicts enhancers with high accuracy and resolution	Success depends on prior knowledge of the main players in the regulatory network of interest	[22,99]
		Relatively low false positives and false negatives, as confirmed through <i>in vivo</i> studies of <i>Drosophila</i> mesoderm development	Relies on availability of multiple ChIP-grade reagents	
Coactivator occupancy-based mapping	ChIP-seq of broadly expressed transcriptional coactivators (e.g. p300/CBP, Mediator)	High accuracy and resolution of mapping	Does not distinguish active and poised enhancers	[14–16,18,35,38]
		Low false positives, as confirmed for p300-based approach via <i>in vivo</i> studies in mice	When relying on a single coactivator (e.g. p300), some enhancers independent of this coactivator will be missed, resulting in false negatives	
Histone modification-based mapping	ChIP-seq profiling of histone modifications (e.g. H3K4me1, H3K27ac, H3K9ac)	ChIP-grade antibodies broadly available and applicable for use across metazoan species	Single modification not informative (e.g. only a fraction of H3K4me1 is found at enhancers)	[17,18,21,24,48,60]
			Accuracy of mapping based solely on histone modifications has not been systematically verified <i>in vivo</i>	
		Robust ChIPs with high DNA recoveries; less starting material can be used compared with ChIPs with TFs or coactivators	Diminished resolution and accuracy of annotation, resulting from the fact that histone modifications show broad peaks and mark nucleosomes that flank, rather than bind, enhancers, often	

Approach	Method	Advantages	Limitations	Refs
			asymmetrically in relation to enhancer position	
Combinatorial mapping	ChIP-seq profiling of a coactivator or RNA Pol II or TFs, combined with analysis of histone modifications patterns (e.g. H3K4me1, H3K27ac, H3K9ac)	High accuracy and resolution of mapping of both enhancer position and activity	Demanding simultaneous enrichments for many factors and/or modifications may increase false negatives	[16,18,19,44]
		Distinguishes active and poised or disengaged enhancers		