Leading Edge
Review
Cell

Looping Back to Leap Forward: Transcription Enters a New Era

Michael Levine,[1,*] Claudia Cattoglio,[1,2] and Robert Tjian[1,2,*]
[1]Department of Molecular and Cell Biology
[2]Howard Hughes Medical Institute, CIRM Center of Excellence, Li Ka Shing Center for Biomedical and Health Sciences
University of California, Berkeley, Berkeley, CA 94707, USA
*Correspondence: mlevine@berkeley.edu (M.L.), jmlim@uclink4.berkeley.edu (R.T.)
http://dx.doi.org/10.1016/j.cell.2014.02.009

Comparative genome analyses reveal that organismal complexity scales not with gene number but with gene regulation. Recent efforts indicate that the human genome likely contains hundreds of thousands of enhancers, with a typical gene embedded in a milieu of tens of enhancers. Proliferation of *cis*-regulatory DNAs is accompanied by increased complexity and functional diversification of transcriptional machineries recognizing distal enhancers and core promoters and by the high-order spatial organization of genetic elements. We review progress in unraveling one of the outstanding mysteries of modern biology: the dynamic communication of remote enhancers with target promoters in the specification of cellular identity.

Introduction

Transcription regulation is the premier mechanism underlying differential gene activity in animal development and disease. The first paradigms of gene control were established in bacteria and phage transcription, which typically employs the promoter as the exclusive site for integrating the information required to switch genes on or off (e.g., Ptashne, 2005). The earliest well-studied systems consisted of a repressor bound to specific ''operator'' sequences that overlapped the promoter, thus precluding entry of RNA polymerase (Pol). Eviction of the repressor by allosteric changes accompanying the binding of an inducer (e.g., lactose), for example, permits access of Pol to the promoter and activation of gene expression (Lewis, 2013). An equally important mechanism regulating Pol binding was revealed by the discovery of sigma factors and activators that help recruit and stabilize Pol at the promoter (Losick, 1998). These also operate in promoter-proximal regions, generally within 50–60 bp of the transcription start site. Thus, in the majority of cases, bacteria, phages, and other prokaryotes rely on promoter-proximal, topologically restricted *cis*-elements to drive regulated transcriptional initiation.

In the late 1970s, scientists obtained the first glimpses into the organization of metazoan genes. When compared with bacteria, three fundamental differences were immediately apparent. First, genes are interrupted by intervening sequences, or introns (Sharp, 1994). Second, the DNA template is wrapped up in nucleosomes, making access to chromatin by *trans*-acting factors a more arduous task (Kornberg and Lorch, 1999). Third, it was possible to identify regulatory DNA sequences—enhancers—extended distances along the DNA from their cognate core promoter. This separation was first dramatically demonstrated in the case of the prototypic enhancer, identified in the animal virus SV40 (Banerji et al., 1981). The entire SV40 genome is only 5.2 kb in length. It contains a 200 bp enhancer located immediately upstream of the early promoter, which controls the expression of genes (e.g., T-antigen) required for replication of the viral genome. The close proximity of the SV40 enhancer to the T-antigen promoter was evocative of the promoter-proximal regulatory elements of bacteria and yeast. However, despite this proximity, the SV40 enhancer was shown to augment the expression of a linked, heterologous gene (β-globin) over a distance of 10 kb, farther than the entirety of the native SV40 genome. This unexpected uncoupling of regulatory DNAs from their target promoters—regulation at a distance—appears to be a distinctive property of metazoan genomes. Although yeast and other simple eukaryotes contain a few genes with such long distance *cis*-control arrangements, the vast majority of their genes employ regulatory sequences located near (100–200 bp) promoters (Struhl et al., 1998). By contrast, the majority of metazoan regulatory DNAs encompass multiple clusters of enhancers located at long distances from their promoters, and recent studies have provided dramatic examples of super long-range enhancer-promoter interactions in vertebrate genomes. For example, the gene encoding Sonic Hedgehog is regulated by a distal enhancer that maps nearly 1 megabase from the promoter (Amano et al., 2009). Moreover, the expression of the *c-Myc* oncogene in hematopoietic lineages is regulated by a cluster of remote enhancers located 1.8 megabases downstream of the transcription unit (Shi et al., 2013). It has been recognized for some time that expanding the tether between the core promoter and *cis*-control elements allows regulation at a distance and opens the door to complex gene control, whereby a given gene can be expressed in a variety of different cell types and tissues and in response to different signals or environmental cues (e.g., Levine, 2010; Bulger and Groudine, 2011). Indeed, we might posit that, without unhitching enhancers and promoters, it would not be

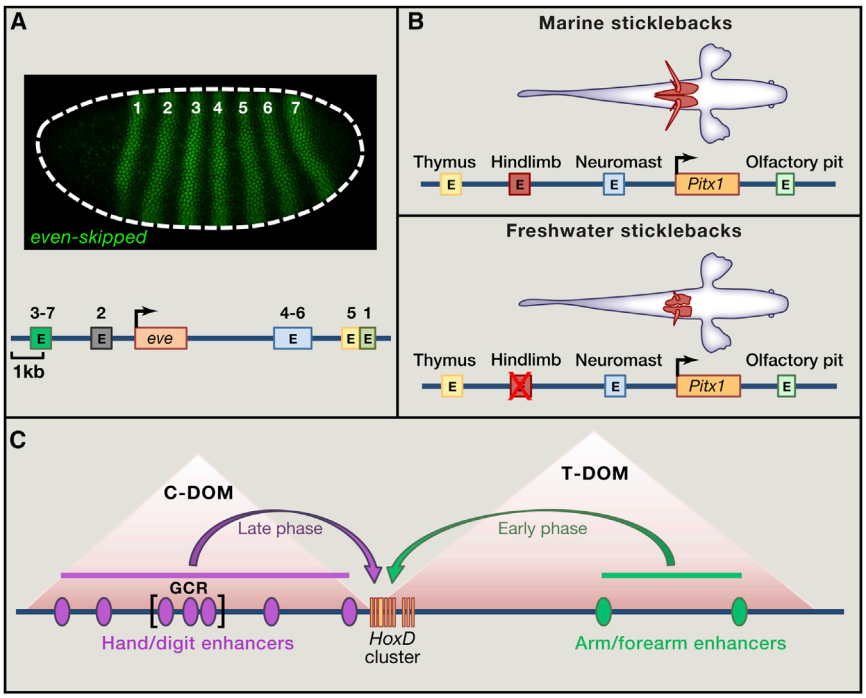

Figure 1. Organization of *cis*-Regulatory DNAs in Metazoan Genomes

Metazoan genes are regulated by multiple enhancers.

(A) Organization of the *even-skipped* (*eve*) locus in the *Drosophila* genome. The *eve* gene is just 3 kb in length but is regulated by individual stripe enhancers (E) located in both 5′ and 3′ flanking regions. The *eve* stripe enhancers function in an additive fashion to produce seven stripes of gene expression in the early *Drosophila* embryo (micrograph by Mike Perry and Michael Levine, personal communication).

(B) Evolution of pelvic fins in stickleback fish. The *Pitx1* gene is regulated in different tissues by a series of enhancers located in both 5′ and 3′ flanking regions. Deletion of the hindlimb enhancer results in reduced development of the pelvic fins (red) in freshwater populations (adapted from Shapiro et al. [2004]).

(C) Organization of the *Hoxd* complex in mice. The complex is regulated by a series of flanking enhancers (purple and green ovals) located in two neighboring TADs. The telomeric TAD (T-DOM) regulates linked *Hoxd* genes in the developing arm and forearm, whereas the centromeric TAD (C-DOM) regulates expression in the hand and the digits (adapted from Andrey et al. [2013]).

possible to assemble the elaborate networks of gene transcription that control complex metazoan processes, and, hence, "location matters" in the evolution of *cis*-regulatory elements. A well-studied example of such a transcription network is seen for the segmentation gene *even-skipped*, which is expressed in seven pair-rule stripes along the length of the *Drosophila* embryo due to the activities of five separate enhancers (Figure 1A) (Levine, 2010). Similarly, the vertebrate *Pitx* gene is regulated by several enhancers mediating expression in different tissues and organs (Chan et al., 2010). Selective deletions of the hindlimb enhancer underlie the diversification of stickleback fish populations lacking pelvic fins (Figure 1B). Thus, the modular organization and distal locations of metazoan enhancers enabled the development of multiple cell types and likely facilitated animal evolutionary diversity. A major future frontier of transcription research is the elucidation of the dynamic communication of remote enhancers with their target promoters. Before delving into this ambitious topic, we first summarize recent advances in our understanding of the protein complexes controlling the activity of RNA polymerase II (Pol II) at the core promoters of protein-coding genes and some noncoding RNAs. We emphasize key findings and concepts obtained during the past 10 years and refer the reader to previous reviews for more in-depth discussions of specialized topics, such as histone modification (e.g., Ruthenburg et al., 2007), in silico identification of enhancers based on clustering of recognition sequences for cell-specific transcription factors (e.g., Philippakis et al., 2006), and mechanisms of transcriptional elongation (e.g., Smith and Shilatifard, 2013). We will also not touch on RNA polymerase I and III that transcribe ribosomal RNAs and tRNAs, respectively, not for lack of interest or importance but primarily for consideration of length and focus.

The Core Promoter and Cell-Specific Transcription Complexes

Compared to bacteria, metazoan systems employ vastly more elaborate *trans*-acting protein machineries to cope with the extended arrangement of distal enhancers and multifaceted promoters and the demands of temporal and spatial patterns of gene transcription essential to governing cell-type specificity and development (Levine and Tjian, 2003). In the past 10–15 years, it has been well documented that animal genomes contain a large proportion of genes encoding transcription factors (5%–10% of total coding capacity). There is also a great diversity and functional specialization of components that make up the core promoter recognition complex and "basal machinery." Instead of a relatively simple Pol complex composed of just five to six subunits, the prototypic eukaryotic preinitiation complex (PIC) consists of >85 polypeptides, including several multisubunit components such as RNA Pol II, TFIID, E, F, H, and various large coactivators (Med/ARC) and chromatin-remodeling factors (Roeder, 1996; Cramer, 2002; Goodrich and Tjian, 2010). In recent years, an even greater diversity of cell-type- and gene-specific cofactors and PIC components has been discovered (Figure 2). We have also come to appreciate that even the core promoter comes in many flavors, with elements such as TATA, INR, and DPE contributing additional levels of specificity and regulation when coupled to upstream and downstream enhancers (Juven-Gershon and Kadonaga, 2010). Pol II consists of 12 subunits, is highly conserved throughout eukaryotes, and serves as the central catalytic component of the PIC that drives RNA synthesis (Roeder, 1996). It does so, however, with the help of a large and diverse set of essential core promoter initiation factors that include TFIIA, B, D, E, F, and H. Although most RNA Pol II initiation complexes utilize all or most of these

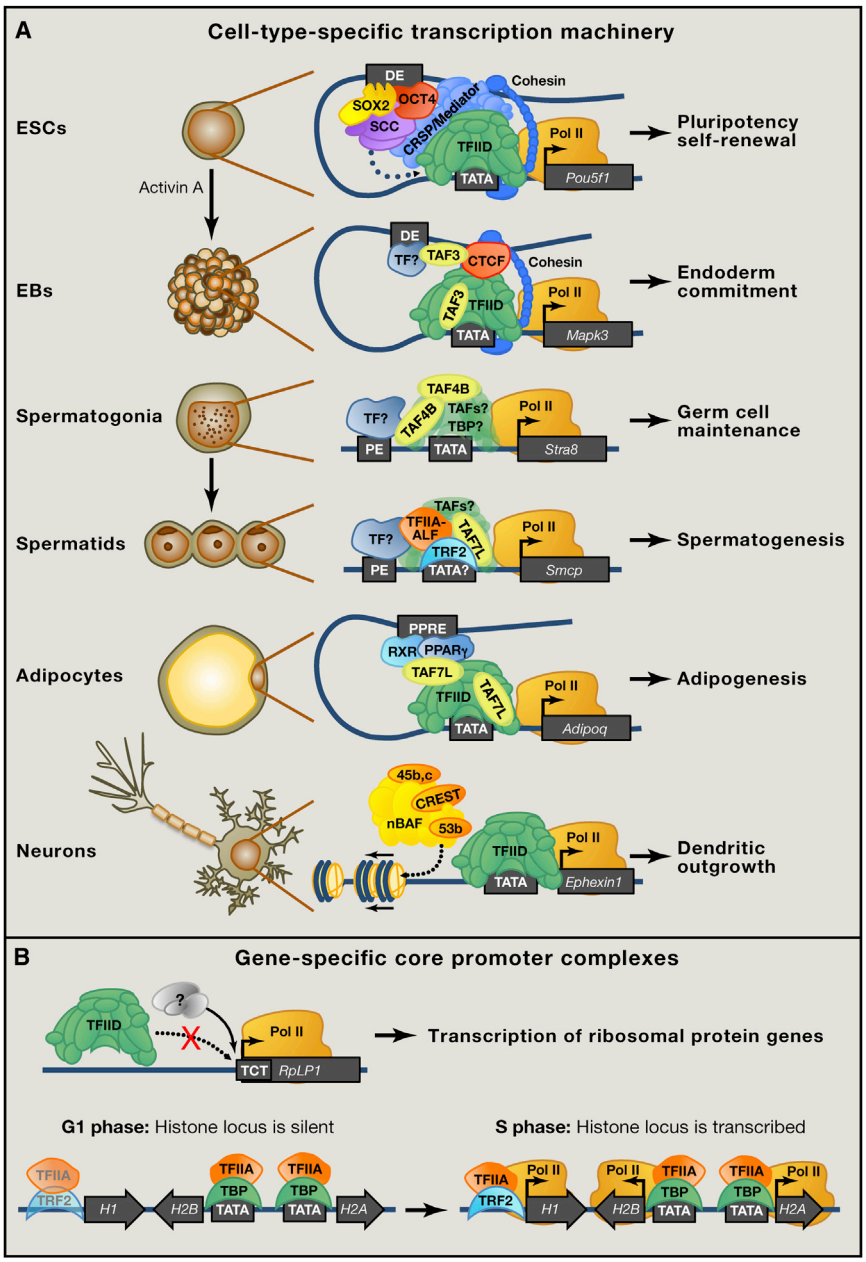

Figure 2. Specialized Transcription Machineries

(A) A diversified set of PICs, coactivators, and chromatin remodelers orchestrates cell-specific transcription programs. In embryonic stem cells (ESCs), the XPC trimeric complex works as an OCT4/SOX2 stem cell coactivator (SCC) at distal enhancer sites (DE) to sustain the expression of pluripotency and self-renewal genes. Upon formation of embryoid bodies (EBs), TBP-associated factor TAF3 is required for endodermal lineage differentiation, mediating DNA looping between DEs and core promoters (TATA) of endoderm-specification genes in concert with CTCF. In testis, TAF4B directs a transcription program required to preserve the germ-cell compartment; farther down the differentiation path, in round spermatids, TAF4B is replaced by a core-promoter complex composed of the TAF7 homolog TAF7L, TBP-related factor TRF2 and TFIIA, which promotes spermatogenesis instead. TAF7L also regulates adipogenesis by associating with TBP as a component of TFIID at promoters and with PPARγ-RXR as a cofactor at enhancers on adipocyte-specific genes. In neurons, a specialized BAF chromatin-remodeling complex exists (nBAF) that includes neural specific subunits (BAF53b, BAF45b, BAF45c, and CREST) and facilitates transcription of genes involved in dendrite outgrowth.

(B) A yet-uncharacterized, TFIID-independent PIC assembles at the TCT motif (polypyrimidine initiator) encompassing the transcription start site of ribosomal protein genes in *Drosophila* cells. In *Drosophila* S2 cells, noncanonical PICs made of TRF2/TFIIA and TBP/TFIIA are responsible for the cell-cycle-restricted expression of H1 and H2B/A histone genes, respectively. TBP/TFIIA, and possibly TRF2/TFIIA, are preloaded on the histone locus in the G1-phase of the cell cycle but only activate transcription when cells enter S phase. Abbreviations: PE, proximal enhancer; PPRE, PPARγ response element; TF, sequence-specific transcription factor.

prototypic core promoter factors (previously referred to as general transcription factors), a wealth of biochemical fractionation studies, functional reconstitution assays, and genetic analyses have revealed that even the core components of the PIC are neither "general" nor universal. It now seems likely that there are various classes of PIC assemblies and that the stereotypic PIC composition identified in human HeLa cells, *Drosophila* S2, and yeast cell extracts may have given us an oversimplified picture that significantly underestimated the diversity of PICs (Goodrich and Tjian, 2010). The importance of the PIC and constituent subunits in specifying cell-type-selective gene regulation became apparent only after mechanisms of transcriptional control were examined in terminally differentiated cells (Deato and Tjian, 2007). It now appears that many components of the PIC, as well as attendant coactivators and chromatin remodeling complexes, come in diverse ensembles that are required to drive cell-type- and gene-specific transcription in metazoans (Figure 2A) (D'Alessio et al., 2009; Dikstein et al., 1996; Goodrich and Tjian, 2010). This notion of diversified and functionally distinct sets of Pol II accessory factors working in concert with classical sequence-specific DNA-binding activators to regulate cell-type-specific transcriptional programs may itself be an underestimate, as it now seems evident that there are also gene specific PICs within a single cell type (Figure 2B). For example, it was recently found that the transcriptional complexes responsible for expression of the histone genes in *Drosophila* make use of a rather stripped down version of the stereotypic PIC—one that lacks both TFIID and B (Guglielmi et al., 2013). Another striking example of gene-specific PICs is seen for the genes encoding ribosomal proteins (Figure 2B), wherein the core promoters bear a distinct TCT element instead of the more conventional

TATA/INR core promoter arrangement (Parry et al., 2010). Thus, despite over 30 years of extensive biochemical and genetic analysis, it is likely that we have not yet completed our survey of the core Pol II machineries that govern metazoan transcription, and we look forward to additional surprises.

Structure of Multisubunit Complexes that Form the PIC

There have been major advances in the determination of the 3D structures of the large multisubunit assemblies that are recruited to the core promoter to form the PIC (Grünberg and Hahn, 2013; Liu et al., 2013). The earliest successes in the structural determination of transcription factors came from elegant X-ray crystallographic studies of sequence-specific DNA-binding proteins (McKay and Steitz, 1981; Pavletich and Pabo, 1991). These studies helped define the now easily recognizable structural motifs such as zinc fingers, helix-loop-helix, leucine zippers, homeobox, winged helix, and many other DNA-binding domains (Weirauch and Hughes, 2011). By contrast, the 3D structures of transcriptional activation domains (e.g., acidic, glutamine rich, etc.) have stubbornly eluded structural determination. The paucity of such structures results from their inherent lack of stable tertiary structure—"molten globules"—until they interact with specific coactivators, which impose a more defined 3D configuration. In several cases, these disordered domains contain low-complexity (LC) sequences, made up of few repeated amino acid residues. These LC domains that are postulated to form reversible fibrous polymers have recently been identified in Mediator subunits, TAF15, and the largest subunit of RNA Pol II (Kwon et al., 2013). Multiple LC-domain-containing transcription factors binding to arrayed promoter elements provide a potential mechanism to seed the sequential assembly of the PIC. For example, the LC, repetitive C-terminal domain (CTD) of RNA Pol II has the ability to interact with fibrous polymers formed by other LC proteins, and phosphorylation of the CTD disrupts this interaction, allowing regulated promoter escape and elongation. These properties of floppy domains assuming an induced fit structure or multimerizing to create docking platforms are likely to apply to other regulatory proteins encoded in the genomes of higher eukaryotes. Unfortunately, most other critical components of the transcriptional apparatus such as TAFs, coactivators, chromatin remodelers, and core-promoter factors do not bear such obvious signature motifs identifiable through their amino acid sequences (Pavlopoulou and Michalopoulos, 2011). This complicates the task of identifying such factors based on primary sequence information. Consequently, discoveries of new cofactors and core promoter components depend on biochemical fractionation and in vitro transcription assays (Pugh and Tjian, 1990; Tjian and Maniatis, 1994). The crystal structure of Pol II provided new insights into the enzymatic mechanisms of transcription, particularly elongation. Structural analysis revealed that the two largest subunits, Rpb1 and Rpb2, form an active central cleft (Cramer et al., 2000) with the Rpb1 side of the cleft, forming a clamp that is open in the absence of template DNA but closed when template DNA and RNA are present (Gnatt et al., 2001). During elongation, the DNA enters the cleft, where it forms a DNA-RNA hybrid in the active center (Gnatt et al., 2001). Structures of the elongation complex also revealed how RNA Pol II selects NTPs and incorporates them into the nascent RNA, and identifies the exit path of the RNA transcript (Cheung et al., 2011; Kettenberger et al., 2004; Wang et al., 2006; Westover et al., 2004). Although less is known about the mechanisms of initiation, several structural studies have highlighted the role of TFIIB, which appears to interact with Pol II in a manner analogous to the way that sigma factors interact with bacterial Pol despite a lack of sequence homologies (Bushnell et al., 2004; Murakami et al., 2002). Of course, the initiation process requires the participation of several other core PIC components, and a more complete mechanistic dissection of transcription will benefit from high-resolution structures of the PIC discussed below. The large sizes of promoter-associated transcription complexes have limited the use of conventional X-ray crystallographic methods for their elucidation. There is the added challenge that these complexes often adopt alternative configurations and conformations on promoter templates. However, recent improvements in high-resolution EM methods have led to some significant advances, including the structural determinations of the human TFIID complex and a nearly complete PIC (Cianfrocco et al., 2013; He et al., 2013; Liu et al., 2009; Murakami et al., 2013). EM structures of human TFIID bound to DNA revealed that this core component of the PIC induces even more dramatic turns of the promoter template than the binding of TBP to TATA DNA (Figure 3A). Moreover, both TFIID and the ARC/CRSP mediator complexes undergo conformational changes when bound to activators (Figures 3B and 3C) (Liu et al., 2009; Taatjes et al., 2002). Collectively, these structural studies paint a picture of highly flexible, conformationally diverse, and multipronged interactions occurring as activators, Pol II, core promoter factors, elongation factors and chromatin remodeling complexes all converge at the promoter to form the PIC and initiate transcription.

Pioneer Factors

We now consider the current state of knowledge regarding the communication of distal enhancers with the ensembles of transcription complexes present at or near the core promoter. One of the most important new insights arising from whole-genome analyses of animal development is that many genes are systematically primed for their timely activation upon receipt of appropriate inducing signals (reviewed by Lagha et al. [2012]). Both distal enhancers and the core promoter anticipate the subsequent activation of gene expression (Figure 4). Studies on the regulation of liver-specific gene activity in mouse embryos led to the identification of FOXA as a "pioneer" transcription factor that primes enhancers for future activation. FOXA is expressed throughout the developing foregut of early mouse embryos (Ang et al., 1993; Monaghan et al., 1993). The liver arises from a subset of these cells prior to their separation from the foregut FOXA "marks" enhancers that will become active during later stages of development (Gualdi et al., 1996). It was suggested that the binding of FOXA to these enhancers primes them for future induction by opening local chromatin and facilitating the entry of liver-specific transcription factors (Cirillo et al., 2002). The exact mechanism is uncertain. Perhaps FOXA recruits a histone-modifying enzyme like CBP/p300 (Visel et al., 2009) or a chromatin-remodeling enzyme like SWI/SNF (Ronan et al., 2013), which, in turn, renders neighboring factor-binding sites

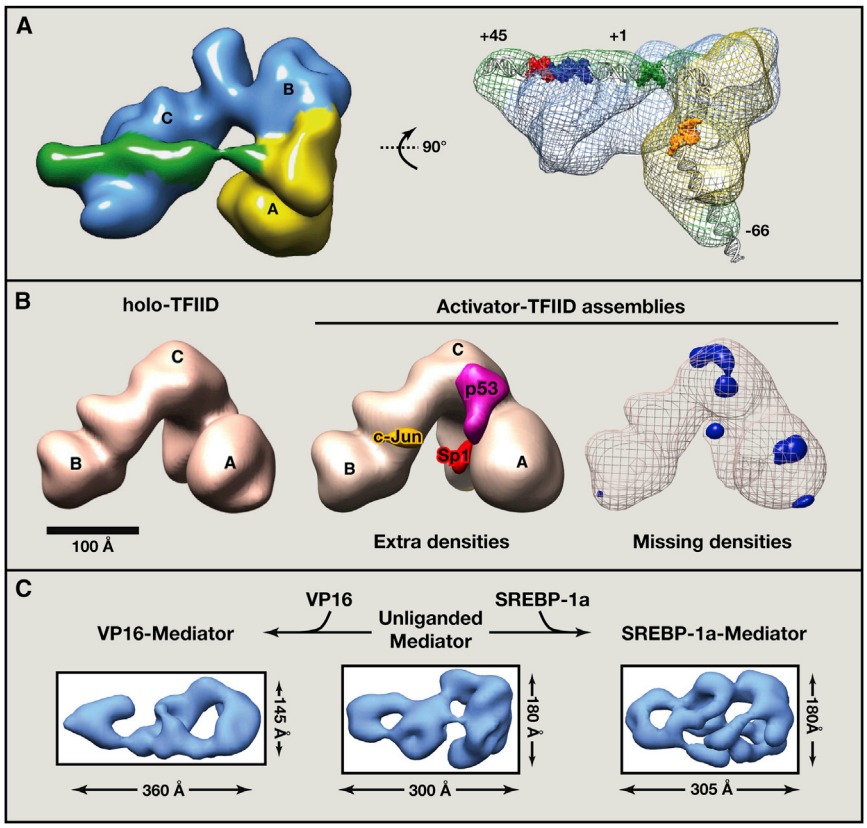

Figure 3. Structural Dynamics of Transcription Machineries

(A) Binding of PIC components to promoter induces dramatic turns of the DNA template, as revealed by EM structure of human TFIID and TFIIA bound to a super-core promoter (adapted from Cianfrocco et al. [2013]).

(B) Different activators (P53, c-JUN, SP1) target distinct sites and induce localized, as well as common, conformational changes within TFIID, as evaluated by EM structural studies (adapted from Liu et al. [2009]).

(C) ARC/CRSP mediator undergoes dramatic and distinct conformational changes when bound to VP16 versus SREBP-1a activators, as resolved by EM (adapted from Taatjes et al. [2002]).

accessible for efficient occupancy once the later liver-specific transcription factors are first expressed. It is now appreciated that transcription factors involved in reprogramming cells from one fate to another, such as OCT4 and SOX2, have pioneer-like properties (Soufi et al., 2012; Wapinski et al., 2013). FOXA exhibits an unusual property during the cell cycle; it remains associated with condensing chromosomes during mitosis, whereas most sequence-specific transcription factors appear to be released (Caravaca et al., 2013). This special quality—avid binding to condensed chromatin—might be a critical manifestation of the ability of pioneer factors to stably mark enhancers for future use (Zaret and Carroll, 2011). Pioneer factors are not a distinctive property of vertebrate systems but appear to occur in invertebrates as well. For example, the *Drosophila* embryo exhibits a sharp transition in gene activation between 2 and 3 hr after fertilization. This is referred to as the maternal-zygotic transition (MZT) and is akin to the midblastula transition in *Xenopus* embryos (Tadros et al., 2007). Early embryogenesis is driven by maternal products deposited into the unfertilized egg during oogenesis. After depletion of these products, the zygotic genome is induced for subsequent developmental processes. A sequence-specific transcription factor called Zelda appears to function as a pioneer factor to mark most or all of the enhancers slated for induction during the MZT (Harrison et al., 2011; Liang et al., 2008; Nien et al., 2011). Zelda is maternally expressed and appears to interact with target enhancers during the first 30–60 min of embryogenesis, well before they become active about an hour later. Altogether, Zelda marks hundreds of enhancers regulating more than 100 zygotic genes. However, unlike FOXA, Zelda has not been specifically shown to remain associated with mitotic chromosomes, which we might expect to see if it functions as a bona fide pioneer factor.

It is likely that the priming of distal enhancers by pioneer (or putative pioneer) factors such as FOXA and Zelda is related to DNaseI hypersensitivity documented for select regulatory sequences such as β-globin genes in classical studies performed in the 1980s (e.g., Groudine et al., 1983). Whole-genome assays suggest that the "marking" of regulatory sequences for future use is a prevalent property of metazoan gene control during development.

Paused RNA Polymerase

In addition to marking their enhancers for future use, developmentally regulated genes can also anticipate activation by acquiring paused Pol II during embryogenesis (Adelman and Lis, 2012; Levine, 2011). Whole-genome Pol II ChIP-Seq and Gro-Seq assays reveal that the majority of genes that contain Zelda at their distal enhancers prior to activation during *Drosophila* MZT also contain paused Pol II (Figure 4C) (Chen et al., 2013; Saunders et al., 2013). Thus, both distal regulatory sequences and the proximal promoter are "primed" for timely induction of gene expression during the 1 hr period, between 2 and 3 hr after fertilization, when localized stripes and bands of gene expression establish the basic blueprint of the adult fly (Levine, 2010). Genes that are activated later in development are not paused during this early time period. Instead, they acquire paused Pol II later, ~1 hr or so prior to their expression (Chen et al., 2013). However, it is important to note that paused Pol II is not the only way to prepare the promoter for timely activation. For example, TBP and TFIIA mark the promoter regions of histone genes throughout the cell cycle, whereas Pol II is recruited only upon transcriptional activation (Figure 2B) (Guglielmi et al., 2013). The mechanism by which paused Pol II (or other general factors such as TFIIA) foster timely activation of gene expression is uncertain. Studies in cultured S2 cells suggest that paused

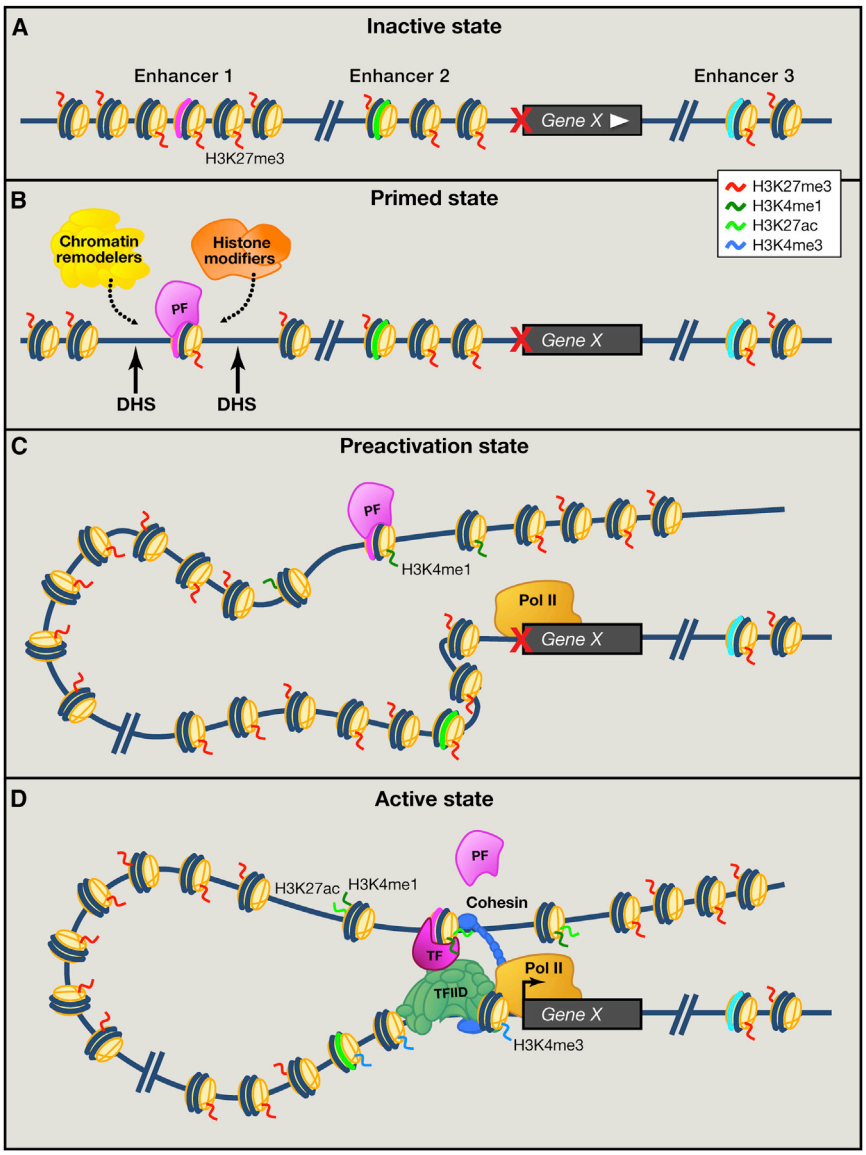

Figure 4. A Model for the Sequential Activation of Gene Expression

Diagram of a hypothetical gene regulated by several distal enhancers located both 5′ and 3′ of the transcription unit.

(A) Gene X is silent; all enhancers are inactive and contain "repressive marks"—H3K27me3—mediated by Polycomb silencers (e.g., Voigt et al., 2013).

(B) A pioneer factor (PF) binds to specific sites in Enhancer 1. This leads to the appearance of flanking DnaseI hypersensitive sites and, presumably, the recruitment of chromatin-remodeling complexes (e.g., BAF) and histone-modifying complexes (e.g., Hu et al., 2013).

(C) Following changes in chromatin state, the regulatory region becomes condensed, thereby bringing Enhancer 1 into proximity with the Gene X promoter. In some cases, the promoter acquires paused Pol II prior to induction.

(D) Upon binding of inductive sequence-specific transcription factors (TF), the Enhancer engages the promoter and leads to the recruitment of the PIC or release of paused Pol II to trigger expression. Cohesin has been implicated in stabilizing enhancer-promoter interactions (e.g., Guo et al., 2012).

Pol II functions very much like pioneer factors, serving as a "bookmark" to prime the gene for future activation (Gilchrist et al., 2010). In particular, it was suggested that paused Pol II blocks the assembly of inhibitory nucleosomes within the core promoter. According to this model, the priming of distal regulatory sequences and proximal promoters coordinates efficient activation upon receipt of appropriate inducing signals (Figure 4). Thus, paused genes are immediately activated upon induction, whereas genes lacking paused Pol II might exhibit stochastic patterns of activation, possibly arising from variable delays in the eviction of inhibitory nucleosomes from the core promoter (Lagha et al., 2013).

Genome-wide Identification of Enhancers

The human genome contains a scant number of genes, on the order of just 25,000, approximately the same number of genes seen in the mustard weed *Arabidopsis* and the nematode worm *C. elegans* (Putnam et al., 2008; Simakov et al., 2013). Given this uncoupling between gene number and organismal complexity, we previously argued that complexity depends on increasingly sophisticated mechanisms of gene regulation (Levine and Tjian, 2003). Thus, the human genome is likely to contain a significantly larger number of enhancers than that seen in plants or worms. Recent studies using whole-genome methods are entirely consistent with this point of view. Several methods have been used for the systematic identification of enhancers, or putative enhancers, engaged in specific developmental processes, including heart specification in mice and the differentiation of cranial neural crest in humans. For example, heart enhancers were identified by examining the genome-wide distribution of CBP/p300 histone acetyltransferase (May et al., 2012). Histone acetylation is associated with transcriptional activation and is seen in both promoter regions and distal enhancers. The idea is that sequence-specific DNA-binding proteins interact with their target sites in distal enhancers and then recruit one or more coactivator complexes to mediate communication with the core promoter. Detection of these complexes, such as CBP/p300, permits identification of active enhancers. These studies suggest that thousands of enhancers control gene expression during the specification and morphogenesis of the mammalian heart. The exact number of authentic enhancers identified by CBP/p300 chromatin immunoprecipitation sequencing (ChIP-seq) assays remains uncertain because only a small fraction of the predicted enhancers was directly tested in transgenic mouse embryos. Another approach to the

whole-genome identification of distal enhancers concerns the use of specific histone modifications, particularly histone H3K4me1 (monomethylation of core histone H3 on lysine 4) and H3K27Ac (acetylation of lysine 27) (Heintzman et al., 2009). These modifications are often associated with latent or active enhancers and have been used to identify distal regulatory DNAs controlling a number of processes, including the specification of cranial neural crest underlying the patterning of the human face (e.g., Calo and Wysocka, 2013). These histone modifications identified thousands of putative neural crest enhancers, and a significant fraction contains sequence polymorphisms in human populations (Rada-Iglesias et al., 2012). Specific polymorphisms were shown to alter the binding of two key sequence-specific transcription factors responsible for the differentiation of cranial neural crest, TFAP2A and NR2F1/2. It is likely that altered binding of such factors underlies some human facial variations and malformations such as cleft palate (Attanasio et al., 2013). The ENCODE consortium identified 400,000 putative enhancers in the human genome, and it is possible that this number could increase to as many as a million enhancers (Bernstein et al., 2012). This amounts to a remarkable fraction of our genomes—25% and probably more—devoted to regulatory information, suggesting that a typical human gene might be regulated by tens of enhancers. Simpler creatures appear to possess fewer enhancers; for example, *Drosophila* appears to contain something like 50,000–100,000 enhancers (Arnold et al., 2013). We emphasize that the current estimates are somewhat uncertain because they are mainly based on whole-genome binding assays, and relatively few direct functional tests. Nonetheless, it is becoming increasingly clear that metazoan genomes are riddled with enhancers.

Enhancer Trafficking and Spatial Organization in the Nucleus

Our view of the regulatory genome has changed significantly in recent years. It now appears that genes are embedded in vast and complex regulatory landscapes. How do the right enhancers communicate with the right promoters in time and space? Chromosome conformation capture methods suggest that mammalian genomes are organized in a series of topological association domains (TADs) composed of an average of 700–800 kb containing 5–10 genes and a few hundred enhancers (Gibcus and Dekker, 2013; Jin et al., 2013). Most enhancer-promoter interactions occur within TADs, although there is evidence for *trans*-TAD and even *trans*-chromosomal interactions. Some TADs are bigger than others. For example, the TAD containing the *c-Myc* locus is ~2 Mb in length and includes a cluster of remote enhancers located at one of the boundaries that controls expression in hematopoietic lineages (Shi et al., 2013). The *Hoxd* complex is located at the cusp of two regulatory TADs, each containing tens of separate enhancers (a "regulatory archipelago"; Figure 1C). The 3′ TAD controls early expression in proximal regions of the developing limb, whereas the 5′ TAD controls expression in the distal regions of the limb that form the digits. The 5′ TAD contains the global control region (GCR), which maps ~200 kb away from the *Hoxd* complex (see de Laat and Duboule, 2013). It is ~40 kb in length and is composed of multiple enhancers. The GCR is evocative of the β-globin LCR, which is responsible for temporal switching of linked globin genes (Martin et al., 1996; see below). Whole-genome analysis of H3K4me1 and H3K27Ac profiles identified putative "super-enhancers" as extended regulatory sequences, spanning ~5–50 kb in length (Whyte et al., 2013). There are more than 200 super-enhancers in the human genome, and most appear to be associated with key regulatory genes specifying particular cell types, such as *Oct4* in embryonic stem cells. It is possible that super-enhancers function much like the previously identified LCR at the β-globin locus and GCR at the *Hoxd* locus, which coordinate the expression of linked genes in defined cell lineages (e.g., Figure 1C). Preliminary studies suggest that a number of disease-associated sequence polymorphisms map within super-enhancers (Hnisz et al., 2013). Indeed, sequence polymorphisms in the noncoding, regulatory genome are emerging as an important source of human variation, including susceptibility to disease. The structured 3D spatial organization of the genome and the manner in which genes and regulatory elements are embedded therein suggest an important role for the relative location of these sequences in 3D space in facilitating regulation of gene expression.

Enhancer-Promoter Communication

One of the outstanding mysteries of transcription regulation is the nature of enhancer-promoter communication. It has been more than 30 years since the discovery of the prototypic SV40 enhancer (Banerji et al., 1981), and yet, we still do not understand the dynamics of this process. There is considerable evidence for the looping of distal enhancers to the promoter. For example, the SP1 activator binds to both enhancers and proximal regions of target promoters, and homotypic interactions between SP1 subunits have been shown to promote and stabilize looping interactions (Dynan and Tjian, 1983; Su et al., 1991). Indeed, a number of promoter-proximal binding proteins do not appear to function as classical activators (e.g., acidic activation domain) but instead might augment gene expression by facilitating communication of distal enhancers with their target promoters (e.g., Calhoun and Levine, 2003). A vivid illustration of the importance of enhancer looping was recently documented at the mouse β-globin gene. As discussed earlier, β-globin is regulated by the looping of the distal LCR to the β-globin promoter (Martin et al., 1996). Looping and activation depend on two key transcription factors, GATA1 and LDB1 (Deng et al., 2012). Both proteins bind to the LCR and β-globin promoter. Removal of GATA1 blocks globin expression. However, expression is restored with a synthetic ZF::LDB1 fusion protein that recognizes specific sequences in the globin promoter and fosters LCR looping in the absence of GATA1. This bypass experiment highlights the importance of enhancer looping in gene activation during development. Recent whole-genome binding assays have identified general transcription factors at distal enhancers. The pregenome view of such factors, e.g., subunits of the Mediator and TFIID complexes, is that they assemble at or near the promoter to foster transcription initiation. However, whole-genome assays have identified binding of TAF3 and TAF7L, so-called "orphan TAFs," at both core promoters and distal enhancers in ES cells and adipocytes, respectively (Figure 2A) (Liu et al., 2011; Zhou et al., 2013). Similar methods have also identified the Pol II elongation factor

ELL3 at distal enhancers (Lin et al., 2013). These observations raise the possibility that the assembly of a fully functional PIC might depend on interactions of distal regulatory sequences with promoter-proximal elements. According to this view, distal enhancers work synergistically with the core promoter to activate transcription.

Integrating Posttranscriptional Processes, DNA Replication, and Repair

Over the past 40 years, increasing evidence points to a coordinated crosstalk between transcription and various steps along the flow of information from DNA replication to protein production, as well as related DNA transactions such as maintenance of genome integrity. There is emerging evidence that DNA repair processes may be employed for the orderly trafficking of the genomic regulatory landscape. For example, BAF/BRG1 and BRM-associated factors (SWI/SNF-like complexes) remodel chromatin at distal enhancers and have been implicated in a variety of development and disease processes. BAF was recently shown to recruit topoisomerase IIα and mediate decatenation of sister chromatids during mitosis (Dykhuizen et al., 2013). It is possible that this topo II activity is also required for long-range enhancer-promoter interactions. There are additional examples of DNA repair enzymes functioning as potential coactivators at distal enhancers. The XPC complex can serve both as a classical DNA repair factor and as a transcriptional coactivator in ES cells (Figure 2A) (Fong et al., 2011). Moreover, components of the nucleotide excision repair pathway have been implicated in transcriptional activation upon DNA demethylation and gene looping (Le May et al., 2012), whereas the base-excision repair enzyme TDG is emerging as a key player in regulating DNA methylation (Wyatt, 2013). Transcription is also coupled to RNA splicing and processing. Key components of transcription initiation and elongation, including CRSP/Mediator (Huang et al., 2012), RNA Pol II, and P-TEFb (Zhou et al., 2012), modulate RNA splicing and processing. Conversely, splicing factors such as SF2/ASF influence transcription levels affecting RNA Pol II pausing and elongation rates (Zhou et al., 2012). Recent studies raise the possibility that specialized splicing byproducts—circular intronic RNAs resistant to debranching—might regulate the expression of their parent genes (Zhang et al., 2013). Despite the physical sequestration of transcription in the nucleus from protein synthesis in the cytoplasm, eukaryotic cells exhibit a surprisingly tight coordination of these processes (Dahan and Choder, 2013). For example, the efficacy of protein synthesis from mRNA templates is determined by specific sequence motifs in 5′ or 3′ UTRs and associated factors that are loaded onto pre-RNAs during transcription (Haimovich et al., 2013a). Examples of such coordination include the yeast RNA Pol II subunits Rpb4/7p (Dahan and Choder, 2013), the Ccr4-Not complex (Miller and Reese, 2012), the human CEBP1 (Bava et al., 2013), and ELAV/Hu (Simone and Keene, 2013) proteins. Likewise, there is evidence that mRNA degradation can affect transcriptional output. In yeast, reduced mRNA decay rates are balanced by diminished mRNA synthesis, so that steady-state mRNA levels are maintained (Sun et al., 2013). Critical for balancing mRNA levels are cytoplasmic decay factors such as Xrn1, which were shown to translocate to the nucleus and work as transcription regulators (Haimovich et al., 2013b). In short, these findings collectively suggest that the linear view of gene expression from DNA to RNA to protein should instead be viewed as circular—transcription affects and is affected by its downstream processes. There is also a mounting body of evidence for the coupling of transcription with DNA replication. The timing of DNA replication in S phase is coupled to transcription levels, whereby genes that are highly expressed are replicated early, while the majority of late-replicating genes are silent (Dellino et al., 2013). Such a temporal correlation between DNA replication and the onset of expression is seen for the large histone gene cluster (Guglielmi et al., 2013). It has been suggested that components of the DNA replication machinery associate with distal regulatory sequences to modulate the timing of transcription (Forsburg, 2004; Karmakar et al., 2010). There is also a spatial component, with respect to where genes are positioned in the nucleus and relative to each other, to the timing of replication from different origins (e.g., Gilbert et al., 2010). Thus, again in the instance of the interplay of transcription and replication, it appears that location within the 3D "nucleome" matters.

Emerging Technologies and the Future of Gene Regulation

During these past several decades, powerful new technologies have been added to the traditional biochemical and genetic methods used to investigate transcription, including an explosion of techniques for genome-wide high-throughput analysis, leading to a "global systems" view of gene regulation (de Wit and de Laat, 2012; van Steensel and Dekker, 2010). Thus, we witnessed an inexorable shift from single-gene analysis to whole-genome surveys, sometimes with interesting and unexpected results. We strongly expect this postgenome era and affiliated systems-level analyses to continue for some time. In addition, a new and equally compelling technology is beginning to have a big impact in the field, especially in revealing new dynamic spatial and temporal aspects of transcription: single-molecule live-cell imaging. A looming future challenge is the direct visualization of enhancer-promoter interactions to determine not only the 3D disposition and relative location of these critical *cis*-regulatory elements but also the time dimension and temporal cadence of their interactions with transcription factors. It is currently unclear how long it takes for a distal enhancer to "find" its target promoter and the stability of the ensuing enhancer-promoter complex. Moreover, it is not known how many rounds of PIC assembly, initiation, and reinitiation result from a single enhancer-promoter interaction. Many of us who were trained in traditional in vitro biochemistry or classic genetic approaches could not imagine the revolution in molecular imaging that has now been sweeping into the life sciences. The idea that we could one day actually see and track the movement of individual transcription factors functioning in living cells in real time or measure reaction dynamics and spatial resolution of single molecules within individual cell nuclei seemed beyond reach. These questions can now be addressed by the powerful new imaging methods that permit the detection of single molecules in living cells and tissues (Figure 5A) (Darzacq et al., 2009; Mueller et al., 2013; Xia et al., 2013). These methods offer the promise of tracking the movements and behaviors of individual

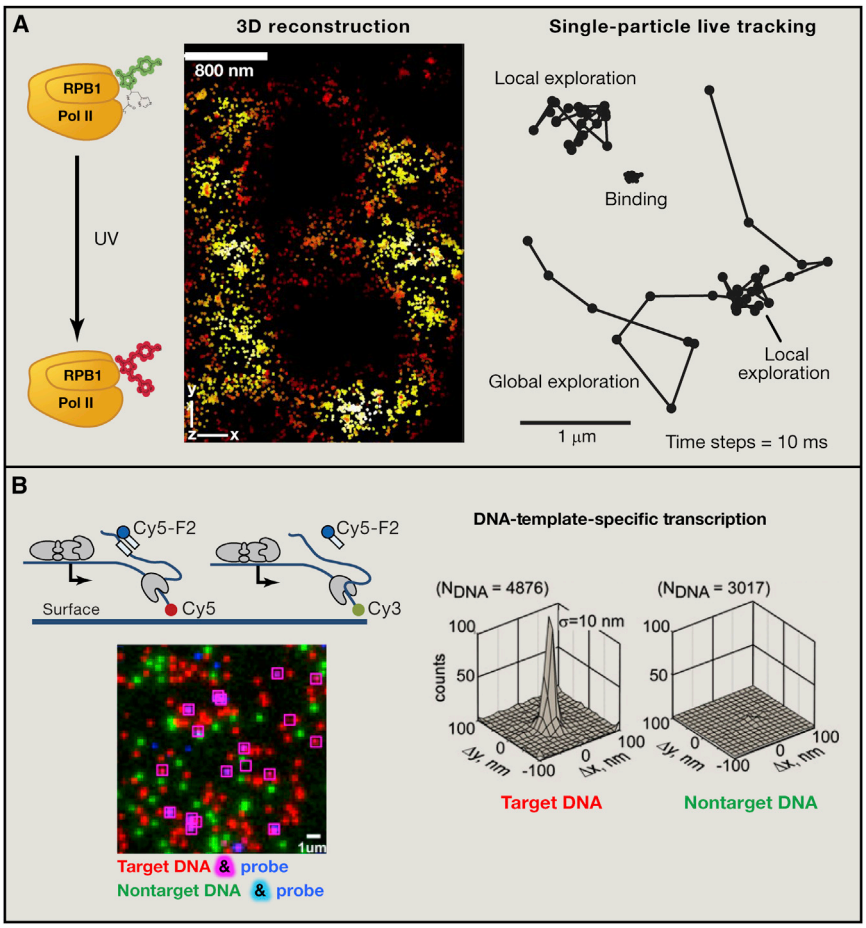

Figure 5. Emerging Imaging Technologies

(A) Single-molecule super-resolution imaging of RNA Pol II (I. Izeddin, I. Cisse, M. Dahan, and X. Darzacq, personal communication). 3D density map of Pol II localization in fixed nuclei (left) highlights spatial Pol II clustering, whereas single-particle tracking in live cells (right) identifies distinct Pol II dynamic behaviors. Data were collected from an engineered cell line stably expressing the Pol II catalytic subunit (RPB1) labeled with the photoconvertible fluorescent protein Dendra2 (Cisse et al., 2013).

(B) Promoter-specific transcription initiation directed by a reconstituted human Pol II system at single-molecule resolution using TIRF video microscopy. Cy5-labeled DNA templates containing a consensus Pol II promoter are immobilized on a surface, and nascent transcripts are detected based on colocalization of fluorescent probes and template signals. The two DNA templates contain (red) or lack (green) the target sequence for the transcript probe to control for specificity (adapted from Revyakin et al. [2012]).

transcription factors as they search for cognate binding sites on interphase chromatin within the nucleus of individual living cells in subsecond real-time measurements (Abrahamsson et al., 2013; Betzig et al., 2006; Gao et al., 2012; Huang et al., 2009; Huisken et al., 2004; Shao et al., 2011; Wu et al., 2013). When combined with genetic manipulations, genome-wide analysis, and in vitro single-molecule assays (Revyakin et al., 2012), these methods can provide extraordinarily quantitative measurements with remarkable temporal and spatial resolution.

It is now possible to accurately measure on/off rates, dwell times, 3D diffusion intervals, and search times for individual transcription factors or combinations of transcription factors. These approaches can also allow us to dissect the in vivo order of events (i.e., which TFs must bind first to a site before others can approach and bind) at enhancers and promoters (Chen et al., 2014). We can also probe how mutations in both the transcription factor proteins and *cis*-regulatory sequence elements will alter the search parameters and binding constants. At the same time, it is possible to manipulate and alter the chromatin/epigenetic state of cells using drugs or mutations to assess the consequences of changing specific chromatin modifications on the transcription search pattern and simultaneously measure the transcriptional output. Just as super-resolution imaging has been a game changer for tracking complex molecular transactions in living cells, a parallel but equally enabling set of advances in single-molecule in vitro biochemistry is revolutionizing our ability to dissect the mechanistic steps involved in cell-free single-molecule transcription assays (Figure 5B) (Bustamante et al., 2011; Deniz et al., 2008; Friedman and Gelles, 2012; Herbert et al., 2008; Revyakin et al., 2012; Treutlein et al., 2012). The development of advanced imaging methods, such as light-sheet microscopy, permits the rapid acquisition of cellular images using a new generation of detectors to record movies over periods of hours in living embryos (Keller et al., 2008). The very first movies of gene regulation are beginning to appear. For example, after nearly 30 years of analyzing fixed preparations of staged *Drosophila* embryos, we can finally watch the dynamic activation of *hunchback* expression by the maternal Bicoid gradient, one of the paradigms of gene control in development (Garcia et al., 2013; Lucas et al., 2013). These studies reported the detection of nascent transcripts produced by the proximal *hunchback* enhancer in living embryos in real time. They reveal incredibly rapid induction of gene expression within a factor of two of the theoretical limit (one Pol II complex loaded every 70–80 bp along the DNA template). These studies also demonstrated that low levels of the Bicoid activator gradient result in all or none expression of the *hunchback* reporter gene in neighboring cells, raising the possibility that activators function in a statistical manner to increase the probability of on or off transcription in the different cells of a population. Previous static methods for the analysis of fixed preparations have been useful for elucidating the spatial control of gene expression (e.g., the borders of segmentation stripes of expression). However, these methods provide limited information about the temporal dynamics of gene expression. The newly available imaging technologies provide the first opportunities for delving into the dimension of time. We anticipate that all of these

technical advances, together with the resurgent interest in mammalian development and stem cell biology, bode well for a rich and productive period. We particularly look forward to new insights into the emerging theme of ''location matters,'' that is, the impact of 3D chromosomal organization and nuclear localization in transcriptional dynamics. The upcoming decade of transcription biology is poised for unprecedented opportunities for discovery.

ACKNOWLEDGMENTS

R.T. and C.C. would like to acknowledge the California Institute for Regenerative Medicine (CIRM) for the Research grant (RB4-06016) to R.T. M.L. is supported by grants from the NIH (GM34431 and GM46638). C.C. has been a postdoctoral fellow of California Institute for Regenerative Medicine (CIRM training program TG2-01164).

REFERENCES

Abrahamsson, S., Chen, J., Hajj, B., Stallinga, S., Katsov, A.Y., Wisniewski, J., Mizuguchi, G., Soule, P., Mueller, F., Dugast Darzacq, C., et al. (2013). Fast multicolor 3D imaging using aberration-corrected multifocus microscopy. Nat. Methods *10*, 60–63.

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet. *13*, 720–731.

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

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

Ang, S.L., Wierda, A., Wong, D., Stevens, K.A., Cascio, S., Rossant, J., and Zaret, K.S. (1993). The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. Development *119*, 1301–1315.

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

Attanasio, C., Nord, A.S., Zhu, Y., Blow, M.J., Li, Z., Liberton, D.K., Morrison, H., Plajzer-Frick, I., Holt, A., Hosseini, R., et al. (2013). Fine tuning of craniofacial morphology by distant-acting enhancers. Science *342*, 1241006.

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

Bava, F.A., Eliscovich, C., Ferreira, P.G., Miñana, B., Ben-Dov, C., Guigó, R., Valcárcel, J., and Méndez, R. (2013). CPEB1 coordinates alternative 3′-UTR formation with translational regulation. Nature *495*, 121–125.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M.; ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science *313*, 1642–1645.

Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327–339.

Bushnell, D.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science *303*, 983–988.

Bustamante, C., Cheng, W., and Mejia, Y.X. (2011). Revisiting the central dogma one molecule at a time. Cell *144*, 480–497.

Calhoun, V.C., and Levine, M. (2003). Long-range enhancer-promoter interactions in the Scr-Antp interval of the Drosophila Antennapedia complex. Proc. Natl. Acad. Sci. USA *100*, 9878–9883.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Mol. Cell *49*, 825–837.

Caravaca, J.M., Donahue, G., Becker, J.S., He, X., Vinson, C., and Zaret, K.S. (2013). Bookmarking by specific and nonspecific binding of FOXA1 pioneer factor to mitotic chromosomes. Genes Dev. *27*, 251–260.

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

Chen, K., Johnston, J., Shao, W., Meier, S., Staber, C., and Zeitlinger, J. (2013). A global change in RNA polymerase II pausing during the Drosophila midblastula transition. Elife *2*, e00861.

Chen, J., Zhang, Z., Li, L., Chen, B., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., et al. (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. Cell. Published online March 13, 2014. http://dx.doi.org/10.1016/j.cell.2014.01.062.

Cheung, A.C., Sainsbury, S., and Cramer, P. (2011). Structural basis of initial RNA polymerase II transcription. EMBO J. *30*, 4755–4763.

Cianfrocco, M.A., Kassavetis, G.A., Grob, P., Fang, J., Juven-Gershon, T., Kadonaga, J.T., and Nogales, E. (2013). Human TFIID binds to core promoter DNA in a reorganized structural state. Cell *152*, 120–131.

Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M., and Zaret, K.S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FOXA) and GATA-4. Mol. Cell *9*, 279–289.

Cisse, I.I., Izeddin, I., Causse, S.Z., Boudarene, L., Senecal, A., Muresan, L., Dugast-Darzacq, C., Hajj, B., Dahan, M., and Darzacq, X. (2013). Real-time dynamics of RNA polymerase II clustering in live human cells. Science *341*, 664–667.

Cramer, P. (2002). Multisubunit RNA polymerases. Curr. Opin. Struct. Biol. *12*, 89–97.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. Science *288*, 640–649.

D'Alessio, J.A., Wright, K.J., and Tjian, R. (2009). Shifting players and paradigms in cell-specific transcription. Mol. Cell *36*, 924–931.

Dahan, N., and Choder, M. (2013). The eukaryotic transcriptional machinery regulates mRNA translation and decay in the cytoplasm. Biochim. Biophys. Acta *1829*, 169–173.

Darzacq, X., Yao, J., Larson, D.R., Causse, S.Z., Bosanac, L., de Turris, V., Ruda, V.M., Lionnet, T., Zenklusen, D., Guglielmi, B., et al. (2009). Imaging transcription in living cells. Annu. Rev. Biophys. *38*, 173–196.

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

de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev. *26*, 11–24.

Deato, M.D., and Tjian, R. (2007). Switching of the core transcription machinery during myogenesis. Genes Dev. *21*, 2137–2149.

Dellino, G.I., Cittaro, D., Piccioni, R., Luzi, L., Banfi, S., Segalla, S., Cesaroni, M., Mendoza-Maldonado, R., Giacca, M., and Pelicci, P.G. (2013). Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing. Genome Res. *23*, 1–11.

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

Deniz, A.A., Mukhopadhyay, S., and Lemke, E.A. (2008). Single-molecule biophysics: at the interface of biology, physics and chemistry. J. R. Soc. Interface *5*, 15–45.

Dikstein, R., Zhou, S., and Tjian, R. (1996). Human TAFII 105 is a cell type-specific TFIID subunit related to hTAFII130. Cell *87*, 137–146.

Dykhuizen, E.C., Hargreaves, D.C., Miller, E.L., Cui, K., Korshunov, A., Kool, M., Pfister, S., Cho, Y.J., Zhao, K., and Crabtree, G.R. (2013). BAF complexes facilitate decatenation of DNA by topoisomerase IIα. Nature *497*, 624–627.

Dynan, W.S., and Tjian, R. (1983). The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell *35*, 79–87.

Fong, Y.W., Inouye, C., Yamaguchi, T., Cattoglio, C., Grubisic, I., and Tjian, R. (2011). A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. Cell *147*, 120–131.

Forsburg, S.L. (2004). Eukaryotic MCM proteins: beyond replication initiation. Microbiol. Mol. Biol. Rev. *68*, 109–131.

Friedman, L.J., and Gelles, J. (2012). Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. Cell *148*, 679–689.

Gao, L., Shao, L., Higgins, C.D., Poulton, J.S., Peifer, M., Davidson, M.W., Wu, X., Goldstein, B., and Betzig, E. (2012). Noninvasive imaging beyond the diffraction limit of 3D dynamics in thickly fluorescent specimens. Cell *151*, 1370–1385.

Garcia, H.G., Tikhonov, M., Lin, A., and Gregor, T. (2013). Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. Curr. Biol. *23*, 2140–2145.

Gibcus, J.H., and Dekker, J. (2013). The hierarchy of the 3D genome. Mol. Cell *49*, 773–782.

Gilbert, D.M., Takebayashi, S.I., Ryba, T., Lu, J., Pope, B.D., Wilson, K.A., and Hiratani, I. (2010). Space and time in the nucleus: developmental control of replication timing and chromosome architecture. Cold Spring Harb. Symp. Quant. Biol. *75*, 143–153.

Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540–551.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876–1882.

Goodrich, J.A., and Tjian, R. (2010). Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. Nat. Rev. Genet. *11*, 549–558.

Groudine, M., Eisenman, R., Gelinas, R., and Weintraub, H. (1983). Developmental aspects of chromatin structure and gene expression. Prog. Clin. Biol. Res. *134*, 159–182.

Grünberg, S., and Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. Trends Biochem. Sci. *38*, 603–611.

Gualdi, R., Bossard, P., Zheng, M., Hamada, Y., Coleman, J.R., and Zaret, K.S. (1996). Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. Genes Dev. *10*, 1670–1682.

Guglielmi, B., La Rochelle, N., and Tjian, R. (2013). Gene-specific transcriptional mechanisms at the histone gene cluster revealed by single-cell imaging. Mol. Cell *51*, 480–492.

Guo, Y., Monahan, K., Wu, H., Gertz, J., Varley, K.E., Li, W., Myers, R.M., Maniatis, T., and Wu, Q. (2012). CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. Proc. Natl. Acad. Sci. USA *109*, 21081–21086.

Haimovich, G., Choder, M., Singer, R.H., and Trcek, T. (2013a). The fate of the messenger is pre-determined: a new model for regulation of gene expression. Biochim. Biophys. Acta *1829*, 643–653.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millán-Zambrano, G., Barkai, O., Chávez, S., Pérez-Ortín, J.E., Darzacq, X., and Choder, M. (2013b). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000–1011.

Harrison, M.M., Li, X.Y., Kaplan, T., Botchan, M.R., and Eisen, M.B. (2011). Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLoS Genet. *7*, e1002266.

He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. Nature *495*, 481–486.

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

Herbert, K.M., Greenleaf, W.J., and Block, S.M. (2008). Single-molecule studies of RNA polymerase: motoring along. Annu. Rev. Biochem. *77*, 149–176.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934–947.

Hu, D., Gao, X., Morgan, M.A., Herz, H.M., Smith, E.R., and Shilatifard, A. (2013). The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Mol. Cell. Biol. *33*, 4745–4754.

Huang, B., Bates, M., and Zhuang, X. (2009). Super-resolution fluorescence microscopy. Annu. Rev. Biochem. *78*, 993–1016.

Huang, Y., Li, W., Yao, X., Lin, Q.J., Yin, J.W., Liang, Y., Heiner, M., Tian, B., Hui, J., and Wang, G. (2012). Mediator complex regulates alternative mRNA processing via the MED23 subunit. Mol. Cell *45*, 459–469.

Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E.H. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science *305*, 1007–1009.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290–294.

Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev. Biol. *339*, 225–229.

Karmakar, S., Mahajan, M.C., Schulz, V., Boyapaty, G., and Weissman, S.M. (2010). A multiprotein complex necessary for both transcription and DNA replication at the β-globin locus. EMBO J. *29*, 3260–3271.

Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science *322*, 1065–1069.

Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol. Cell *16*, 955–965.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285–294.

Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L., and McKnight, S.L. (2013). Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell *155*, 1049–1060.

Lagha, M., Bothma, J.P., and Levine, M. (2012). Mechanisms of transcriptional precision in animal development. Trends Genet. *28*, 409–416.

Lagha, M., Bothma, J.P., Esposito, E., Ng, S., Stefanik, L., Tsui, C., Johnston, J., Chen, K., Gilmour, D.S., Zeitlinger, J., and Levine, M.S. (2013). Paused Pol II coordinates tissue morphogenesis in the Drosophila embryo. Cell *153*, 976–987.

Le May, N., Fradin, D., Iltis, I., Bougnères, P., and Egly, J.M. (2012). XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. Mol. Cell *47*, 622–632.

Levine, M. (2010). Transcriptional enhancers in animal development and evolution. Curr. Biol. *20*, R754–R763.

Levine, M. (2011). Paused RNA polymerase II as a developmental checkpoint. Cell *145*, 502–511.

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature *424*, 147–151.

Lewis, M. (2013). Allostery and the lac Operon. J. Mol. Biol. *425*, 2309–2316.

Liang, H.L., Nien, C.Y., Liu, H.Y., Metzstein, M.M., Kirov, N., and Rushlow, C. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature *456*, 400–403.

Lin, C., Garruss, A.S., Luo, Z., Guo, F., and Shilatifard, A. (2013). The RNA Pol II elongation factor Ell3 marks enhancers in ES cells and primes future gene activation. Cell *152*, 144–156.

Liu, W.L., Coleman, R.A., Ma, E., Grob, P., Yang, J.L., Zhang, Y., Dailey, G., Nogales, E., and Tjian, R. (2009). Structures of three distinct activator-TFIID complexes. Genes Dev. *23*, 1510–1521.

Liu, Z., Scannell, D.R., Eisen, M.B., and Tjian, R. (2011). Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. Cell *146*, 720–731.

Liu, X., Bushnell, D.A., and Kornberg, R.D. (2013). RNA polymerase II transcription: structure and mechanism. Biochim. Biophys. Acta *1829*, 2–8.

Losick, R. (1998). Summary: three decades after sigma. Cold Spring Harb. Symp. Quant. Biol. *63*, 653–666.

Lucas, T., Ferraro, T., Roelens, B., De Las Heras Chanes, J., Walczak, A.M., Coppey, M., and Dostatni, N. (2013). Live imaging of bicoid-dependent transcription in Drosophila embryos. Curr. Biol. *23*, 2135–2139.

Martin, D.I., Fiering, S., and Groudine, M. (1996). Regulation of beta-globin gene expression: straightening out the locus. Curr. Opin. Genet. Dev. *6*, 488–495.

May, D., Blow, M.J., Kaplan, T., McCulley, D.J., Jensen, B.C., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., et al. (2012). Large-scale discovery of enhancers from human heart tissue. Nat. Genet. *44*, 89–93.

McKay, D.B., and Steitz, T.A. (1981). Structure of catabolite gene activator protein at 2.9 A resolution suggests binding to left-handed B-DNA. Nature *290*, 744–749.

Miller, J.E., and Reese, J.C. (2012). Ccr4-Not complex: the control freak of eukaryotic cells. Crit. Rev. Biochem. Mol. Biol. *47*, 315–333.

Monaghan, A.P., Kaestner, K.H., Grau, E., and Schütz, G. (1993). Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. Development *119*, 567–578.

Mueller, F., Stasevich, T.J., Mazza, D., and McNally, J.G. (2013). Quantifying transcription factor kinetics: at work or at play? Crit. Rev. Biochem. Mol. Biol. *48*, 492–514.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science *296*, 1285–1290.

Murakami, K., Elmlund, H., Kalisman, N., Bushnell, D.A., Adams, C.M., Azubel, M., Elmlund, D., Levi-Kalisman, Y., Liu, X., Levitt, M., et al. (2013). Architecture of an RNA polymerase II transcription pre-initiation complex. Science *342*, 1238724.

Nien, C.Y., Liang, H.L., Butcher, S., Sun, Y., Fu, S., Gocha, T., Kirov, N., Manak, J.R., and Rushlow, C. (2011). Temporal coordination of gene networks by Zelda in the early Drosophila embryo. PLoS Genet. *7*, e1002339.

Parry, T.J., Theisen, J.W., Hsu, J.Y., Wang, Y.L., Corcoran, D.L., Eustice, M., Ohler, U., and Kadonaga, J.T. (2010). The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. Genes Dev. *24*, 2013–2018.

Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809–817.

Pavlopoulou, A., and Michalopoulos, I. (2011). State-of-the-art bioinformatics protein structure prediction tools (Review). Int. J. Mol. Med. *28*, 295–310.

Philippakis, A.A., Busser, B.W., Gisselbrecht, S.S., He, F.S., Estrada, B., Michelson, A.M., and Bulyk, M.L. (2006). Expression-guided in silico evaluation of candidate cis regulatory codes for Drosophila muscle founder cells. PLoS Comput. Biol. *2*, e53.

Ptashne, M. (2005). Regulation of transcription: from lambda to eukaryotes. Trends Biochem. Sci. *30*, 275–279.

Pugh, B.F., and Tjian, R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell *61*, 1187–1197.

Putnam, N.H., Butts, T., Ferrier, D.E., Furlong, R.F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J.K., et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. Nature *453*, 1064–1071.

Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. Cell Stem Cell *11*, 633–648.

Revyakin, A., Zhang, Z., Coleman, R.A., Li, Y., Inouye, C., Lucas, J.K., Park, S.R., Chu, S., and Tjian, R. (2012). Transcription initiation by human RNA polymerase II visualized at single-molecule resolution. Genes Dev. *26*, 1691–1702.

Roeder, R.G. (1996). Nuclear RNA polymerases: role of general initiation factors and cofactors in eukaryotic transcription. Methods Enzymol. *273*, 165–171.

Ronan, J.L., Wu, W., and Crabtree, G.R. (2013). From neural development to cognition: unexpected roles for chromatin. Nat. Rev. Genet. *14*, 347–359.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. Nat. Rev. Mol. Cell Biol. *8*, 983–994.

Saunders, A., Core, L.J., Sutcliffe, C., Lis, J.T., and Ashe, H.L. (2013). Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription. Genes Dev. *27*, 1146–1158.

Shao, L., Kner, P., Rego, E.H., and Gustafsson, M.G. (2011). Super-resolution 3D microscopy of live whole cells using structured illumination. Nat. Methods *8*, 1044–1046.

Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jónsson, B., Schluter, D., and Kingsley, D.M. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature *428*, 717–723.

Sharp, P.A. (1994). Split genes and RNA splicing. Cell *77*, 805–815.

Shi, J., Whyte, W.A., Zepeda-Mendoza, C.J., Milazzo, J.P., Shen, C., Roe, J.S., Minder, J.L., Mercan, F., Wang, E., Eckersley-Maslin, M.A., et al. (2013). Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. Genes Dev. *27*, 2648–2662.

Simakov, O., Marletaz, F., Cho, S.J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D.H., Larsson, T., Lv, J., Arendt, D., et al. (2013). Insights into bilaterian evolution from three spiralian genomes. Nature *493*, 526–531.

Simone, L.E., and Keene, J.D. (2013). Mechanisms coordinating ELAV/Hu mRNA regulons. Curr. Opin. Genet. Dev. *23*, 35–43.

Smith, E., and Shilatifard, A. (2013). Transcriptional elongation checkpoint control in development and disease. Genes Dev. *27*, 1079–1088.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994–1004.

Struhl, K., Kadosh, D., Keaveney, M., Kuras, L., and Moqtaderi, Z. (1998). Activation and repression mechanisms in yeast. Cold Spring Harb. Symp. Quant. Biol. *63*, 413–421.

Su, W., Jackson, S., Tjian, R., and Echols, H. (1991). DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. Genes Dev. *5*, 820–826.

Sun, M., Schwalb, B., Pirkl, N., Maier, K.C., Schenk, A., Failmezger, H., Tresch, A., and Cramer, P. (2013). Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. Mol. Cell *52*, 52–62.

Taatjes, D.J., Näär, A.M., Andel, F., 3rd, Nogales, E., and Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. Science *295*, 1058–1062.

Tadros, W., Westwood, J.T., and Lipshitz, H.D. (2007). The mother-to-child transition. Dev. Cell *12*, 847–849.

Tjian, R., and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. Cell *77*, 5–8.

Treutlein, B., Muschielok, A., Andrecka, J., Jawhari, A., Buchen, C., Kostrewa, D., Hög, F., Cramer, P., and Michaelis, J. (2012). Dynamic architecture of a minimal RNA polymerase II open promoter complex. Mol. Cell *46*, 136–146.

van Steensel, B., and Dekker, J. (2010). Genomics tools for unraveling chromosome architecture. Nat. Biotechnol. *28*, 1089–1095.

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854–858.

Voigt, P., Tee, W.W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes Dev. *27*, 1318–1338.

Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., and Kornberg, R.D. (2006). Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. Cell *127*, 941–954.

Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F., et al. (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell *155*, 621–635.

Weirauch, M.T., and Hughes, T.R. (2011). A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution. Subcell. Biochem. *52*, 25–73.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell *119*, 481–489.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Wu, Y., Christensen, R., Colón-Ramos, D., and Shroff, H. (2013). Advanced optical imaging techniques for neurodevelopment. Curr. Opin. Neurobiol. *23*, 1090–1097.

Wyatt, M.D. (2013). Advances in understanding the coupling of DNA base modifying enzymes to processes involving base excision repair. Adv. Cancer Res. *119*, 63–106.

Xia, T., Li, N., and Fang, X. (2013). Single-molecule fluorescence imaging in living cells. Annu. Rev. Phys. Chem. *64*, 459–480.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev. *25*, 2227–2241.

Zhang, Y., Zhang, X.O., Chen, T., Xiang, J.F., Yin, Q.F., Xing, Y.H., Zhu, S., Yang, L., and Chen, L.L. (2013). Circular intronic long noncoding RNAs. Mol. Cell *51*, 792–806.

Zhou, Q., Li, T., and Price, D.H. (2012). RNA polymerase II elongation control. Annu. Rev. Biochem. *81*, 119–143.

Zhou, H., Kaplan, T., Li, Y., Grubisic, I., Zhang, Z., Wang, P.J., Eisen, M.B., and Tjian, R. (2013). Dual functions of TAF7L in adipocyte differentiation. Elife *2*, e00170.