Mechanisms of gene activation by metazoan enhancers: New insights and opportunities from high-throughput approaches

Beagrie, R.A., Pombo, A.

Mechanisms of gene activation by metazoan enhancers

New insights and opportunities from high-throughput approaches

Robert A. Beagrie[1*], Ana Pombo[1*]

[*]Corresponding authors:
Robert A. Beagrie
E-mail: Robert.Beagrie@mdc-berlin.de

Ana Pombo
E-mail: Ana.Pombo@mdc-berlin.de

Addresses:

[1]Epigenetic Regulation and Chromatin Architecture Group, Berlin Institute for Medical Systems Biology, Max-Delbrück Centre for Molecular Medicine, Robert-Rössle Strasse, Berlin-Buch 13092, Germany.

Abstract

The molecular characteristics of enhancers are increasingly well defined, and single-gene studies have proposed many different mechanisms to explain how these regulatory DNA sequences act to stimulate the transcription of their target genes. However, which of these mechanisms are general modes of enhancer action remains unknown. Important new insights have been driven by recent developments in high-throughput sequencing, massively parallel reporter gene monitoring and CRISPR-Cas9 technologies. The opportunities presented by these approaches pave the way for the efficient translation of sequence variants affecting enhancers into effective medical treatments.

Keywords:

Enhancer, gene regulation, transcription factors, mediator, eRNA, chromatin conformation.

Introduction: Enhancers are important contributors to human disease

Enhancers are DNA sequences that are capable of activating transcription of a target gene when linked in *cis* to the promoter (i.e. on the same DNA molecule), and do so independently of their orientation with respect to the target gene [1]. Enhancers are now thought to be widespread in mammalian genomes, with estimates ranging from 50 thousand to 500 thousand total enhancers in human cells [2, 3]. Importantly, many genome sequence variants that have been linked to human disease overlap putative enhancers [4, 5]. Therefore, mutations which alter the behaviour of enhancers, and lead to abnormal gene expression patterns rather than malfunctioning gene products, might be a dominant cause of human disease.

In spite of their importance, the mechanisms by which enhancers activate the expression of their target genes are poorly understood, and most of our knowledge is derived from painstaking studies of single genes. In the first half of this review, we summarise the various proposed models of enhancer action. In the second half, we outline recently developed high-throughput methods for assaying enhancer function and highlight areas of opportunity where the application of these strategies could yield the greatest advances in understanding gene regulation in health and disease.

Mechanisms of enhancer action

Transactivation by enhancer-bound transcription factors

In the transactivation model, RNA polymerase (pol) II binds to the promoter along with other components of the pre-initiation complex (PIC; Box 1). Transcription is then stimulated by direct physical interactions with transcription factors (TFs) bound to the enhancer (Fig. 1A). TF "effector domains" activate the target gene in a number of ways, e.g. by stabilising the PIC, increasing the rate of productive initiation by RNA pol II or by enhancing the rate of transcriptional elongation [6]. Protein structural studies have been invaluable for unpicking the interactions between effector domains and the basal transcriptional machinery (especially the Mediator complex; Box 1).

Metazoan enhancers generally comprise a cluster of binding sites for different TFs [7], which can be separated from their target genes by large distances of up to 1 Mb [8]. Much of the

research into TF effector domain function has focused on single TFs bound directly to gene promoters, so it is unclear how far the principles established in these pioneer studies will apply to transactivation by distal TF clusters (Fig. 1A). For example, binding sites for the TFs YY1 and NF-E2 have been shown to be necessary for full activity of the human β-globin locus control region (LCR; a multi-enhancer element that is able to activate expression of nearby genes independent of its chromosomal position). Each TF is capable of activating transcription when bound directly to promoters, yet binding sites for either TF alone have no activity from more distal positions, indicating that cooperativity between different TFs may be particularly important at distant enhancers [9, 10].

Genomic distance between TF binding sites and their target promoters clearly has a strong influence on TF effector domain function, although only a handful of studies have examined these effects directly [11–14]. For example, the E2F1 transactivation domain can stimulate transcription from the mouse dihydrofolate reductase promoter *in vivo* but the effect is completely abrogated by relocation of the E2F binding site from 66 to 441 bp upstream of the promoter [12]. Similarly, transcriptional activation by the VP16 effector domain can be strongly reduced by an insertion of only 54 bp between the promoter and the effector binding site [11]. Importantly, different effector domains display contrasting abilities to activate transcription from distal versus proximal positions [15], and individual TFs generally favour either promoter-proximal or promoter-distal binding positions *in vivo* [16]. This suggests that TFs bound at distal enhancers might not activate their target genes by the same mechanisms as TFs bound at promoters (Fig. 1A). Effector domains are very poorly conserved across the hundreds of metazoan TFs, which further complicates identification of families of related effector domains that might function similarly [6].

Box 1: Transcriptional initiation and the pre-initiation complex

When added to naked DNA *in vitro*, RNA pol II initiates transcription in an essentially random manner and requires the presence of other co-factors for specific and directional initiation at promoters [17]. These additional protein complexes are the general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. The GTFs, together with RNA pol II, form a pre-initiation complex (PIC) which is capable of locating and binding to

promoters [18]. DNA is then unwound by the helicase activity of TFIIH. RNA pol II undergoes several rounds of abortive initiation, producing short 8-10 nucleotide transcripts, before escaping from the promoter, breaking most of its contacts with the GTFs, and transitioning to elongation [19].

Transcription is regulated not only by the activity of promoter-bound GTFs but also by DNA-binding transactivating factors that regulate development, cell lineage commitment or respond to stimuli. These factors stimulate RNA pol II transcription above basal levels, but only in the presence of the Mediator complex [20]. Mediator was originally isolated from yeast and was shown to be required for activation of an *in vitro* transcription system by GCN4 and GAL4 transcription factors [20]. Subsequent work has revealed that mammalian Mediator consists of 26 subunits plus an additional complex of four proteins including the kinase CDK8 [21]. The range of Mediator's known activities currently include stabilising TFIID binding to DNA, recruiting TFIIE and TFIIF to the pre-initiation complex, stimulating TFIIH mediated phosphorylation of the RNA pol II C-terminal domain and promoting processive elongation [22]. Mediator has therefore been proposed as the primary integrator of activating signals from enhancers [21].

Hit-and-run mechanisms: opening chromatin

Transcription factors interact with a wide array of co-factors and co-activators in addition to the basal transcriptional machinery. Many TFs are known to interact with chromatin remodelling complexes and/or histone modifying enzymes such as p300 [6, 23, 24]. The hit-and-run model proposes that looping interactions between enhancers and genes allow enhancer-bound chromatin modifying enzymes to deposit activating histone marks and/or remodel nucleosomes at target promoters, thereby establishing an open chromatin state that promotes PIC formation (Fig. 1B). Most models for TF transactivation generally postulate a direct or indirect interaction between enhancer-bound proteins and RNA pol II during at least one part of the transcription cycle. In contrast in the "hit-and-run" model, enhancer-promoter pairing occurs prior to promoter activation, possibly transiently.

Three studies have targeted epigenetic modifications to specific enhancer regions by fusing histone modifying enzymes to DNA binding factors. The hit-and-run model predicts that

tethering a chromatin modifying enzyme at a distal enhancer should result in chromatin re-organization at target promoters. Recruitment of the histone acetyltransferase p300 to a β-globin enhancer led to acetylation of H3K27 at the haemoglobin epsilon and gamma genes (~11 and 30kb distant from the enhancer, respectively) [25]. Two other studies targeted a repressive chromatin remodeller, LSD1 (which demethylates H3K4), to enhancer regions. Whilst both reported repression of transcription at nearby genes, the first did not study histone modification levels at target gene promoters [26] and the second found no change in either H3K4me2 or H3K27ac at a target promoter [27]. In all three studies, chromatin was much more efficiently modified at the enhancer sequence itself (where the modifying enzyme was tethered). This makes it difficult to unambiguously attribute epigenetic changes at target gene promoters to direct, long-range activities of the enzyme tethered to distal enhancer regions. Alternatively, opening of chromatin at the enhancer might regulate TF recruitment and transcriptional transactivation upstream of epigenetic changes at the target gene promoter.

Other activation pathways could also function through a hit-and-run mechanism. Distal enhancers are occupied by both chromatin remodelling enzymes capable of destabilising nucleosomes [28, 29] and by Tet family proteins, which can hydroxylate methylated DNA [30]. Therefore, transient enhancer-promoter interactions might allow remodellers to create nucleosome-free regions at target promoters, or allow Tet proteins to remove repressive DNA methylation at promoter CpG islands.

Hit-and-run mechanisms: Polycomb eviction from promoters by distal enhancers

Rather than depositing activating histone marks at a target promoter, hit-and-run enhancers could also function by removing repressive marks. The Polycomb eviction model posits that distal enhancer elements remove repressive Polycomb protein complexes from the promoters of developmental genes [31]. Polycomb complexes are present at the α-globin gene in early development and are normally cleared from the promoter prior to transcriptional activation in mature erythroid cells. Mature cells lacking a key α-globin enhancer display elevated Polycomb occupancy at the α-globin gene, indicating that the enhancer element could be involved in removal of Polycomb proteins from α-globin during normal development [31]. A similar mechanism might operate for the neural gene *Meis2*, where a tissue specific enhancer necessary for midbrain expression of *Meis2* interacts with the Polycomb-repressed *Meis2* promoter. This interaction is formed just before Polycomb is lost from the promoter and

Meis2 begins to be expressed, suggesting that the enhancer-promoter loop might actively clear Polycomb complexes from the *Meis2* promoter [32].

The interplay between transcriptionally active chromatin domains (marked by H3K27ac) and repressive chromatin domains (marked by H3K27me3) is particularly important for the regulation of the *Hox* gene clusters [33]. The four mammalian *Hox* clusters each comprise 9-11 genes encoding key TFs crucial for development of the body plan. Within each cluster, these genes are expressed sequentially (i.e. reflecting their genomic organization) both in space (along body axes) and in time (during development), a phenomenon termed collinearity. Collinear expression of *Hox* genes along the anterior to posterior body axis is accompanied by a progressive shift in the boundary between an active chromatin domain and a repressive domain [34]. Collinear expression of *Hoxd* genes along the developing limb requires the action of distal enhancer sequences located on both sides of the *Hoxd* cluster [35, 36], and involves both a shift in boundary and a switch from repressive to active domain (and vice versa [36]). The distal enhancers are presumably involved in clearing Polycomb from the repressive domain, although whether their involvement is direct or indirect, and whether it is separable from their gene activating roles remains to be determined.

Exactly how enhancers might clear Polycomb proteins from gene promoters is also unclear. Interestingly, the H3K27 demethylase JMJD3 is recruited to fibroblast enhancers by P53 following DNA damage [37]. If demethylases such as JMJD3 can remove H3K27me3 marks at promoters by looping interactions from distal binding positions, this might provide a possible mechanism to disrupt Polycomb inhibition at target genes [38].

The pause-release model

After initiation, RNA pol II "pauses" just downstream of the promoter to different extents, and is released from the paused state into productive elongation through the action of P-TEFb, among other factors (reviewed in [39]). One recent paper has described "anti-pause" enhancers in HEK293 cells and suggested that they can control the transition of RNA pol II from the paused to the elongating state from distal positions [40]. The authors identify a protein-protein interaction between the histone demethylase JMJD6 and BRD4, a known regulator of RNA pol II pause-release via P-TEFb. Knock-down of either interacting partner increases RNA pol II promoter-proximal pausing and decreases transcription of a subset of genes. BRD4 and JMJD6 co-bound to intergenic "anti-pause" enhancers and several anti-

pause enhancers were found to contact promoters over a distance of 5-10kb [40]. BRD4 has been localized to enhancers genome-wide in a number of systems [41–43], suggesting that distal enhancers may be general regulators of RNA pol II pausing.

Other interacting partners of the pausing complex P-TEFb also bind to distal enhancers, lending further support to the pause-release model of enhancer function. The effector domain of the transcription factor MYC can recruit Cyclin T1, a component of P-TEFb and key pausing regulator [44]. MYC binds to both promoters and distal enhancers [45], and therefore could conceivably act to regulate RNA pol II pausing from a distance. Another P-TEFb partner Elongin 3 (Ell3) binds enhancers in mouse embryonic stem cells (mESCs) and affects both RNA pol II occupancy and transcription of nearby genes [46]. Long-range pause-release enhancers might also be involved in the response of endothelial cells to VEGF treatment [47].

The RNA pol II transfer model

Enhancers can directly recruit RNA pol II and other components of the PIC [48–50]. Genome-wide ChIP-seq experiments have confirmed RNA pol II presence at enhancers in a variety of tissues, often prior to activation of target genes [51–53], and it has been proposed that enhancer-bound RNA pol II might be directly transferred to target promoters to activate their expression (Fig. 1D and [50]). Treatment with a pause-release inhibitor of pTEFB (flavopiridol) decreases occupancy of RNA pol II at the *PSA* gene, but increases occupancy at the upstream enhancer [54]. This suggests that blocking pause-release might prevent transfer of the polymerase from enhancer to promoter. Similarly, placement of an insulator element (a sequence capable of blocking enhancer activation) between an enhancer and its target promoter [55], or binding of a zinc-finger binding domain to the promoter (thus blocking RNA pol II elongation) [56], can also increase RNA pol II occupancy at an enhancer whilst decreasing promoter occupancy. RNA pol II transfer has also been observed between the β-globin LCR and β-globin gene *in vitro* [50]. Although RNA pol II transfer is consistent with the genome-wide presence of RNA pol II at distal enhancer elements, it so far lacks support as a general mechanism of enhancer function.

Tracking model for enhancer function

An alternative process by which RNA pol II could be delivered from enhancer to promoter involves RNA pol II “tracking”. In this model, RNA pol II binds to the enhancer and

traverses intervening DNA until it reaches the promoter (Fig. 1E; reviewed in [57, 58]). This mechanism has been proposed for activation of the embryonic ε-globin gene at the human β-globin locus [59] and for activation of the *PSA* gene in response to hormone [60]. Unstable RNA transcripts have been detected from the intervening region between some enhancer-promoter pairs, suggesting that the tracking RNA pol II is transcriptionally active [61–63]. Transcription of the target gene is attenuated by insertion of an RNA pol II terminator between the hGH LCR and the hGH promoter [63], or insertion of another gene between the β-globin gene and its LCR [64], lending further support to a possible pol II tracking model at these loci. Tracking and looping may not be mutually exclusive, as looping interactions between enhancer and promoter have been detected alongside RNA pol II tracking [60].

An intermediate between looping and RNA pol II tracking is the "facilitated tracking" model, in which activating proteins bind to an enhancer and are delivered to a target promoter via a loop that gradually expands from the enhancer to the promoter [65]. This model has been used to explain the crosslinking of enhancer-bound TFs to regions between an enhancer and promoter prior to activation [66]. Interestingly, a similar model has recently been proposed for the formation of topologically associated domains, large chromatin regions that preferentially contact themselves [67, 68]. In the loop extrusion model, chromatin loops are randomly formed and enlarged across the genome, but are constrained by boundary elements [69]. If an enhancer were positioned at such a boundary element, it could employ loop extrusion to "scan" surrounding chromatin for a cognate promoter (or vice versa). Whether enhancers are preferentially positioned at TAD boundaries is an open question, as some studies have reported an enrichment for H3K4me1 (a histone mark predictive of enhancers) at TAD boundaries [70, 71], whilst others have not [67].

Tracking is currently disfavoured as a general model due to the large genomic separations that have been observed between many enhancers and their target promoters [8]. Nevertheless, many enhancers are found close enough to their target genes to make use of a tracking mechanism, which might therefore still represent a major enhancer mechanism [58].

Activation of target promoters by eRNAs

Some enhancers are divergently transcribed to produce enhancer RNAs (eRNAs) [51]. By targeting specific eRNAs for degradation (e.g. through siRNA treatment), several studies have reported that eRNA depletion reduces the transcription of nearby genes [72–79]. Levels

of repressive histone marks (e.g. H3K9me3 or H3K27me3) and/or total histone levels did not increase after siRNA treatment, suggesting that the observed effect on transcription at nearby genes was not a consequence of reduced transcription or decreased accessibility of chromatin at the enhancer element, but was instead indicative of direct functional roles for the eRNA transcripts themselves. Experiments so far support two possible cis-mechanisms. In the first, eRNAs interact directly with promoter bound components (Fig. 2A), and either recruit activating factors like Mediator [74] (similar to the transactivation model) or displace suppressive factors such as NELF [77] (similar to the pause-release model). In the second, eRNAs function as a "trap", serving to increase the affinity of TFs with dual DNA/RNA binding ability for actively transcribed enhancer sequences (Fig. 2B and [80]).

One important consideration is to distinguish eRNAs that act in cis, on local chromatin only (Fig. 2A,B), from trans-acting eRNAs that can act via free diffusion (Fig. 2C). Cis and trans mechanisms can be distinguished using DNA- and RNA-binding fusion proteins, which tether RNAs containing a specific stem-loop structure to genomic locations containing their DNA-binding motif [72, 75]. When expressed from a plasmid, a cis-acting eRNA will only activate its target genes in the presence of a tethering protein, whereas if activation occurs in the absence of tethering, the eRNA likely acts in trans. The eRNA upstream of the MyoD gene is a good example, as it can increase transcription of MyoD when overexpressed from a plasmid [76]. Similarly, the yeast GAL10 lncRNA activates the neighbouring GAL1 gene even when moved to a distant genomic location [81]. Where trans mechanisms are dominant, non-coding transcripts are likely to be more stable [82] and might be better referred to as lncRNAs rather than eRNAs.

If expression from a plasmid is not sufficient for activation, the eRNAs in question must somehow remain localized at their sites of transcription. However, when the local concentration of eRNAs was increased locally on chromatin through depletion of the Integrator RNA-processing complex that interferes with eRNA termination, the transcription of nearby genes was reduced rather than increased [83]. It is therefore unclear whether transcription itself is the only anchor required to retain eRNAs at their locus of origin, or whether an additional mechanism is necessary for localization of eRNAs.

In short, the causal relationship between enhancer transcription and enhancer function is difficult to dissect. In some cases, transcripts are clearly dispensable for gene activation [84], in which case the presence of RNA pol II at the enhancer may be crucial for regulatory

function, or may simply be a by-product of open chromatin combined with physical proximity to RNA pol II at the target promoter. In other cases, the transcription of the enhancer may be important to maintain an open chromatin state whilst the transcripts themselves are dispensable [85]. Multiple mechanisms might be employed by a single element; for example, an enhancer whose DNA sequence alone can activate target gene transcription might also act as the promoter of a lncRNA that has additional activating effects in trans.

Enhancers as regulators of chromatin architecture

One alternative model is that enhancers function as tethering points and are recruited to nuclear regions with a high concentration of activating factors [86] (e.g. transcription factories or splicing speckles). Because enhancers are physically linked on the same chromosome, the recruitment of the enhancer could pull nearby genes into the same nuclear neighbourhood, thereby increasing the local concentration of the transcriptional or RNA-processing machinery around the promoter.

Since promoters might themselves be recruited to active neighbourhoods, this model allows for a functional equivalence between enhancers and promoters. Indeed, enhancers can replace promoters to drive stable transcription of a promoterless luciferase gene [54], and weak promoters can act as enhancers by activating transcription of luciferase in cis in an orientation-independent fashion [87]. It has recently been suggested that enhancers and promoters might be a single class of regulatory element with the activity of any given region determined by its local context [88], which would fit well with a model in which enhancers increase the retention of their target genes in "active" nuclear neighbourhoods.

Distinguishing between models

Many different mechanistic models have been proposed to explain the activation of promoters by distal regulatory enhancers acting at the level of DNA, RNA or both (Fig. 3). Although some models are very similar to one another, they all make different predictions about key aspects of enhancer function (Table 1). Further research will be crucial to disentangle which models are general and which are specific to a small number of genes and their enhancers.

One potential complication is that none of the proposed models are mutually exclusive. Individual enhancers composed of many different binding sites might recruit multiple different TFs and co-factors that each makes use of a different mechanism to activate the expression of the target promoter. Importantly, clusters of different TF binding sites (heterotypic clusters) are common in metazoan genomes [89–91] and can function as enhancers [92]. Synthetic enhancers that bind multiple different TFs generally drive higher levels of reporter expression than enhancers containing the same number of sites all bound by the same TF [93]. Different enhancer mechanisms affect different stages of the transcription cycle (Fig. 3), and enhancers that utilize multiple pathways by binding multiple different TFs could achieve synergistic activation of the target promoter [94]. Dissecting the contributing mechanisms of long-range regulation is a major current challenge in deciphering mechanisms of disease associated with non-coding genetic mutations.

Table 1: Putative mechanisms of distal enhancer action.

Model	Chief support	Testable predictions
Transactivation	Structural studies of TF effector domains in complex with basal transcription machinery	Amino acid substitutions in TFs or PIC components that are predicted to disrupt their interaction interface also abrogate enhancer activity from distal locations
Hit-and-run	Presence of chromatin modifying enzymes at enhancers genome-wide	Enhancer-promoter contacts precede promoter activation; promoter activation is dependent on chromatin modifiers
Pause-release	Knock-down of enhancer bound proteins can lead to RNA pol II pausing at promoters	Promoters with high or low levels of RNA pol II pausing respond to different enhancers; depleting pause-promoting factors (e.g. NELF) can eliminate the activating effect of certain enhancers
RNA pol II transfer	Observation of RNA pol II transfer *in vitro*	Individual RNA pol II molecules bind to enhancers before binding to promoters *in vivo*; enhancers are spatially proximal prior to RNA pol II transfer; preventing enhancer-promoter proximity should diminish transfer as a function of distance; RNA pol II retention at enhancers should prevent recruitment to promoters.

Model	Chief support	Testable predictions
RNA pol II tracking	Detection of unstable transcripts between enhancers and target genes	Enhancers and promoters must be oriented in the same direction. Poly(A) sites inserted downstream of the enhancer block gene activation; enhancer-promoter pairs do not need to be spatially close for RNA pol II transfer to promoter, but time of transfer should depend on linear genomic distance between enhancer and promoter.
Loop extrusion	Chromatin extrusion model can predict effects of CTCF site manipulation	Enhancer-promoter interactions are dependent on the activity of a chromatin translocating "motor"; one partner of each enhancer-promoter pair must be positioned near a properly oriented CTCF site.
Polycomb eviction	Enhancer deletion can lead to Polycomb presence at target promoters	Some enhancers only activate Polycomb repressed promoters; Polycomb clearance precedes promoter activation; enhancer function is dependent on the catalytic activity of histone demethylases/de-ubiquitinases
cis-acting eRNA	Tethering of eRNAs at target promoters can lead to activation	eRNA target genes are dependent on the genomic location of the eRNA; promoter activation is dependent on eRNA, but not DNA sequence
Relocation model	Parsimonious explanation of enhancer, insulator and silencer functions	Enhancer function is dependent on nuclear localization, and could be disrupted by altering tethering to nuclear landmarks

New assays for enhancer function

Generalizing the mechanisms of enhancer function will require a shift from detailed, single-gene studies, to high-throughput, massively parallel assays able to test predictions at hundreds or thousands of genes/enhancers at once. Technologies capable of driving this shift have emerged over the past five years, and have the potential to dramatically improve our understanding of distal enhancer function.

High-throughput reporter assays

One recently developed class of technique are high-throughput reporter assays that can measure the activity of hundreds to thousands of putative enhancer elements in a single experiment (Table 2; reviewed in [95–97]). These techniques can be categorized according to three different variables.

First, some methods are designed to test synthetic DNA fragments of defined sequence [98], whereas other methods test fragments isolated from source genomic DNA [99]. Although synthetic approaches allow for very precise variation of the position, orientation and spacing of TF binding sites within enhancers, they are more expensive and limit the size of the fragment that can be tested.

Second, high-throughput reporter assays can vary in the reporter used. In some cases, the enhancer drives expression of a unique barcode [100], which has the advantage of allowing the measurement of the same enhancer coupled to multiple different barcodes, but generally requires fragment synthesis (see above). Alternatively, the enhancer can drive its own transcription, allowing the identity of active enhancer elements to be determined directly by sequencing the RNAs produced [99]. This approach can be less expensive and less time-consuming, but it is generally less quantitative and cannot exclude that the enhancer sequence affects RNA-stability. Another solution is to FACS-sort cells based on expression of a fluorescent reporter, and measure the enrichment of each enhancer sequence in the high-fluorescence compared to the low-fluorescence population. FACS-based methods can be highly quantitative [101], but they can only be applied to cell types that are amenable to sorting.

Third and finally, many techniques test the activity of enhancers during transient transfection. This is the simplest and highest-throughput approach, but non-integrated plasmids may not replicate the full complexity of gene regulation on chromosomes. For this reason, some groups developed approaches where the reporter/enhancer construct is integrated into the genome either at random locations [102] or at a predefined "neutral" location [103]. Random integration is problematic, as the influence of the local chromatin structure at the integration site may be orders of magnitude greater than the influence of the enhancer sequence on the transcription of the reporter gene [104]. Integration at a defined genomic location addresses these issues, but remains difficult to achieve with great efficiency, and therefore limits the number of sequences that can be assayed in parallel.

The application of these high-throughput reporter assays has already yielded a number of important insights:

1. TF consensus motifs can be used to predict enhancer variants that affect expression [100, 101, 105];
2. Multiple binding sites for the same TF generally increase activity, and for some TFs single binding sites are insufficient for enhancer activity [93, 106];
3. The activity of identical TF binding sites can be highly dependent on local sequence context [106];
4. The activity of TF binding site clusters can be influenced by both spacing and orientation [93, 101, 105];

Massively parallel reporter assays have therefore made significant contributions to our understanding of how enhancers are constructed from constituent TF binding sites. However, creative approaches to reporter design that combine these approaches with other novel technologies will be required to fully exploit their potential and to dissect mechanistic questions about how exactly those TFs regulate the expression of target genes.

Table 2: High-throughput strategies for assaying enhancer activity

Technique	Source of tested sequences	Expression readout	Cellular location	Refs
Synthetic saturation mutagenesis	Programmable microarray	Barcode sequence reads	*in vitro*	[107]
Massively parallel reporter assay (MPRA)	Programmable microarray	Barcode sequence reads	Episomal	[93, 105, 106]
Massively parallel functional dissection (MPFD)	Polymerase cycling assembly	Barcode sequence reads	Episomal	[98]
FACS-based approach	Programmable microarray	Enrichment of barcodes in FACS-sorted population	Episomal	[101]
Cis-regulatory element analysis by sequencing (CRE-seq)	Programmable microarray	Barcode sequence reads	Episomal	[100]
Self-transcribing active regulatory region sequencing (STARR-seq)	Fragmented genomic DNA	Enhancer drives its own expression and is sequenced directly	Episomal	[99]
Enhancer-FACS-seq (eFS-seq)	Individually cloned	Enrichment of cis-regulatory element in FACS-sorted population	Site-specific genomic integration	[108]
Thousands of reporters integrated in parallel (TRIP)	Same sequence - used to test local chromatin effects	Barcode sequence reads	Random genomic integration	[104]

Technique	Source of tested sequences	Expression readout	Cellular location	Refs
Functional identification of regulatory elements within accessible chromatin (FIREWACh)	Enzymatically accessible, nucleosome-free genomic DNA	Enrichment of cis-regulatory element in FACS-sorted population	Random genomic integration	[102]
Site-specific integration FACS followed by sequencing (SIF-seq)	Fragmented bacterial artificial chromosomes (BACs)	Enrichment of cis-regulatory element in FACS-sorted population	Site-specific genomic integration	[103]
Capture STARR-seq (CapStarr-seq)	Captured natural variants from patient population	Enhancer drives its own expression and is sequenced directly	Episomal	[109]
***in situ* saturating mutagenesis**	CRISPR/Cas9 mediated random mutagenesis of endogenous element	Enrichment of cis-regulatory element in FACS-sorted population	Endogenous location	[110]

High-throughput activator bypass assays

Another useful technique for studying TF effector domains is the activator bypass assay [111], where an effector domain is artificially recruited to a normally "inert" DNA sequence. The effect of this recruitment on the activity of a reporter gene can then quantified and separated from local chromatin effects. Studies utilising this technique have been particularly important for our understanding of transactivation [11, 12].

Steady improvements in cloning technology have improved the scalability of these assays. One recent study coupled the approach with massively parallel luciferase assays. TFs were recruited to reporter plasmids using a GAL4 DNA-binding domain, which allowed the effect of 474 different *Drosophila* TFs to be measured in multiple sequence contexts [112]. Local sequence context was found to be extremely important for TF function, and the authors were able to show that TFs can be functionally grouped according to their context-dependency.

Activator bypass studies have generally examined the recruitment of TFs to promoter-proximal sites, and have not studied effects on distal enhancers. As the techniques have become more affordable, it is now easier to address long-range effects. One study in yeast fused 223 chromatin regulator proteins to an engineered zinc-finger binding domain and individually recruited them to a genomically integrated reporter gene [113]. Interestingly, they found that some chromatin regulators have very different effects depending on whether they are tethered upstream or downstream of the reporter. Clearly, there is still a great deal of scope for the use of activator bypass strategies to dissect the effects of local chromatin environment and genomic distance on the regulation of endogenous genes.

Applications for CRISPR-Cas9

The key principle of CRISPR-Cas9 technology is that a single guide RNA (sgRNA) complementary to a given genomic region can target Cas9 nuclease to that location and induce a DNA double strand break. CRISPR-Cas9 has been used to screen large numbers of putative enhancers for activity in their endogenous context [114]. The first high-throughput reporter assays measured the effect of all possible nucleotide substitutions on an enhancer *in vitro* [107], but tiling libraries of sgRNAs can now be used to perform saturating mutagenesis directly on endogenous enhancers *in vivo* [110]. Instead of random mutagenesis, multiplex homology-directed repair can be used to replace an endogenous sequence with thousands of

synthetically designed variants [115]. In the future, this approach could be used to test the effect of systematically deleting, inverting or moving single TF binding sites on endogenous enhancers.

The development of catalytically inactive Cas9 (dCas9) has begun to transform CRISPR into a general technology for recruiting molecules of interest to endogenous genomic loci *in vivo*, as dCas9 fusion proteins can be targeted to specific loci without inducing double strand breaks. Fusing dCas9 to fluorescent proteins has been successfully used to advance imaging of transcription [116] and the positions of specific DNA loci [117]. The technology can be further improved by extending the sgRNA so that it protrudes from the Cas9 complex and exposes one or more RNA hairpins, which can then be bound by fluorescent proteins fused to RNA-binding proteins, allowing multi-colour imaging of different loci [118]. These approaches are compatible with live-cell imaging, and may therefore provide the temporal resolution necessary to distinguish hit-and-run mechanisms for enhancer action from the more widely accepted transactivation model.

Cas9 can also be fused directly to TFs, providing an effective way to increase the throughput of activator bypass assays, and allowing the recruitment of single effector domains [119] or combinations of effectors [120] to endogenous loci or integrated reporter genes. sgRNAs can be extended by fusion with endogenous lncRNAs, effectively tethering these RNAs to specific loci [121], providing a tool to help distinguish cis-acting and trans-acting eRNAs. Importantly, all CRISPR-Cas9 systems greatly advance our ability to examine genomic distance effects, as sgRNAs can be easily replaced to perform the same experimental manipulation at multiple different genomic locations.

Further combining CRISPR-Cas9 based approaches with massively parallel reporter assays could enable even greater experimental throughput. For example, an sgRNA library could be used to tether a given TF at thousands of locations surrounding an integrated reporter gene. Loci that function as enhancers after the addition of a single TF could then be identified by sorting cells based on expression of the reporter and determining which sgRNAs were enriched in the high-expression population. Such combined approaches have enormous potential and could transform our understanding of how enhancers regulate the expression of their target genes over the next decade.

Conclusions

Since the discovery of enhancers over 30 years ago, many models have been proposed to explain how enhancers activate transcription of their target genes. The lack of understanding about which of these are general models impedes the ability to predict how disease-linked non-coding sequence variants will affect the expression of surrounding genes. The proposed mechanisms of enhancer action make different predictions about the activities of proteins involved in transcription, and about the likely dynamics and specificity of enhancer-promoter interactions. New technologies including CRISPR-Cas9 tethering and live-cell imaging of chromatin dynamics will be critically important for testing these predictions, and will pave the way for new treatments based upon enhancer biology.

Abbreviations

3C, chromosome conformation capture; **CRISPR**, Clustered regularly interspaced short palindromic repeats; **eRNA**, enhancer RNA; **FACS**, fluorescence-activated cell sorting; **LCR**, locus control region; **lncRNA**, long non-coding RNA; **mESC**, mouse embryonic stem cell; **PIC**, pre-initiation complex; **pol**, polymerase; **sgRNA**, single guide RNA; **TAD**, topologically associating domain; **TF**, transcription factor.

Acknowledgements

We thank Niall Dillon (MRC-Clinical Sciences Centre, UK) and Markus Schueler (Berlin Institute for Medical Systems Biology, Germany) for comments on the manuscript. The Pombo laboratory is funded by the Helmholtz Foundation, the Berlin Institute of Health and the National Institutes of Health (1U54DK107977-01).

References

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

2. **Pennacchio LA, Bickmore W, Dean A, Nobrega MA,** et al. 2013. Enhancers: five essential questions. *Nat. Rev. Genet.* **14**: 288–95.

3. **Andersson R, Gebhard C, Miguel-Escalada I, Hoof I,** et al. 2014. An atlas of active enhancers across human cell types and tissues. *Nature* **507**: 455–61.

4. **Corradin O**, **Scacheri PC**. 2014. Enhancer variants: evaluating functions in common disease. *Genome Med.* **6**: 85.

5. **Hindorff LA, Sethupathy P, Junkins HA, Ramos EM,** et al. 2009. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci.* **106**: 9362–7.

6. **Frietze S**, **Farnham PJ**. 2011. Transcription Factor Effector Domains. In *A Handbook of Transcription Factors*. p 261–77.

7. **Spitz F**, **Furlong EEM**. 2012. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**: 613–26.

8. **Lettice LA, Horikoshi T, Heaney SJH, van Baren MJ,** et al. 2002. Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci.* **99**: 7548–53.

9. **Philipsen S**, **Talbot D**, **Fraser P**, **Grosveld F**. 1990. The beta-globin dominant control region: hypersensitive site 2. *EMBO J.* **9**: 2159–67.

10. **Talbot D**, **Philipsen S**, **Fraser P**, **Grosveld F**. 1990. Detailed analysis of the site 3 region of the human beta-globin dominant control region. *EMBO J.* **9**: 2169–77.

11. **Roberts SG, Choy B, Walker SS, Lin YS,** et al. 1995. A role for activator-mediated TFIIB recruitment in diverse aspects of transcriptional regulation. *Curr. Biol.* **5**: 508–16.

12. **Fry CJ**, **Slansky JE**, **Farnham PJ**. 1997. Position-dependent transcriptional regulation of the murine dihydrofolate reductase promoter by the E2F transactivation domain. *Mol. Cell. Biol.* **17**: 1966–76.

13. **Dobi KC**, **Winston F**. 2007. Analysis of Transcriptional Activation at a Distance in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **27**: 5575–86.

14. **Reavey CT, Hickman MJ, Dobi KC, Botstein D,** et al. 2015. Analysis of Polygenic Mutants Suggests a Role for Mediator in Regulating Transcriptional Activation Distance in Saccharomyces cerevisiae. *Genetics* **201**: 599–612.

15. **Seipel K**, **Georgiev O**, **Schaffner W**. 1992. Different activation domains stimulate transcription from remote (’enhancer') and proximal ('promoter') positions. *EMBO J.* **11**: 4961–8.

16. **Cheng Y, Ma Z, Kim B-H, Wu W,** et al. 2014. Principles of regulatory information conservation between mouse and human. *Nature* **515**: 371–5.

17. **Weil PA**, **Luse DS**, **Segall J**, **Roeder RG**. 1979. Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* **18**: 469–84.

18. **Thomas MC**, **Chiang C-M**. 2006. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**: 105–78.

19. **Saunders A**, **Core LJ**, **Lis JT**. 2006. Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.* **7**: 557–67.

20. **Kim YJ, Björklund S, Li Y, Sayre MH,** et al. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77**: 599–608.

21. **Allen BL**, **Taatjes DJ**. 2015. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**: 155–66.

22. **Conaway RC**, **Conaway JW**. 2011. Origins and activity of the Mediator complex. *Semin. Cell Dev. Biol.* **22**: 729–34.

23. **Visel A, Blow MJ, Li Z, Zhang T,** et al. 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**: 854–8.

24. **Heintzman ND, Stuart RK, Hon G, Fu Y,** et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**: 311–8.

25. **Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI,** et al. 2015. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**: 510–7.

26. **Mendenhall EM, Williamson KE, Reyon D, Zou JY,** et al. 2013. Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol.* **31**: 1133–6.

27. **Kearns NA, Pham H, Tabak B, Genga RM,** et al. 2015. Functional annotation of native enhancers with a Cas9–histone demethylase fusion. *Nat. Methods* **12**: 401–3.

28. **Yu Y, Chen Y, Kim B, Wang H,** et al. 2013. Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell* **152**: 248–61.

29. **Shi J, Whyte WA, Zepeda-Mendoza CJ, Milazzo JP,** et al. 2013. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev.* **27**: 2648–62.

30. **Pulakanti K, Pinello L, Stelloh C, Blinka S,** et al. 2013. Enhancer transcribed RNAs arise from hypomethylated, Tet-occupied genomic regions. *Epigenetics* **8**: 1303–20.

31. **Vernimmen D, Lynch MD, De Gobbi M, Garrick D,** et al. 2011. Polycomb eviction as a new distant enhancer function. *Genes Dev.* **25**: 1583–8.

32. **Kondo T, Isono K, Kondo K, Endo TA,** et al. 2014. Polycomb Potentiates Meis2 Activation in Midbrain by Mediating Interaction of the Promoter with a Tissue-Specific Enhancer. *Dev. Cell* **28**: 94–101.

33. **Montavon T**, **Duboule D**. 2013. Chromatin organization and global regulation of Hox gene clusters. *Philos. Trans. R. Soc. B Biol. Sci.* **368**: 20120367.

34. **Noordermeer D, Leleu M, Splinter E, Rougemont J,** et al. 2011. The Dynamic Architecture of Hox Gene Clusters. *Science* **334**: 222–5.

35. **Montavon T, Soshnikova N, Mascrez B, Joye E,** et al. 2011. A Regulatory Archipelago Controls Hox Genes Transcription in Digits. *Cell* **147**: 1132–45.

36. **Andrey G, Montavon T, Mascrez B, Gonzalez F,** et al. 2013. A Switch Between Topological Domains Underlies HoxD Genes Collinearity in Mouse Limbs. *Science* **340**: 1234167.

37. **Williams K, Christensen J, Rappsilber J, Nielsen AL,** et al. 2014. The histone lysine demethylase JMJD3/KDM6B is recruited to p53 bound promoters and enhancer elements in a p53 dependent manner. *PLoS One* **9**: e96545.

38. **Kondo T**, **Ito S**, **Koseki H**. 2016. Polycomb in Transcriptional Phase Transition of Developmental Genes. *Trends Biochem. Sci.* **41**: 9–19.

39. **Adelman K**, **Lis JT**. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**: 720–31.

40. **Liu W, Ma Q, Wong K, Li W,** et al. 2013. Brd4 and JMJD6-Associated Anti-Pause Enhancers in Regulation of Transcriptional Pause Release. *Cell* **155**: 1581–95.

41. **Lovén J, Hoke HA, Lin CY, Lau A,** et al. 2013. Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. *Cell* **153**: 320–34.

42. **Zhang W, Prakash C, Sum C, Gong Y,** et al. 2012. Bromodomain-containing protein 4 (BRD4) regulates RNA polymerase II serine 2 phosphorylation in human CD4+ T cells. *J. Biol. Chem.* **287**: 43137–55.

43. **Anand P, Brown JD, Lin CY, Qi J,** et al. 2013. BET bromodomains mediate transcriptional pause release in heart failure. *Cell* **154**: 569–82.

44. **Eberhardy SR**, **Farnham PJ**. 2002. Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. *J. Biol. Chem.* **277**: 40156–62.

45. **Sabò A, Kress TR, Pelizzola M, de Pretis S,** et al. 2014. Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature* **511**: 488–

92.

46. **Lin C, Garruss AS, Luo Z, Guo F,** et al. 2013. The RNA Pol II elongation factor Ell3 marks enhancers in ES cells and primes future gene activation. *Cell* **152**: 144–56.

47. **Kaikkonen MU, Niskanen H, Romanoski CE, Kansanen E,** et al. 2014. Control of VEGF-A transcriptional programs by pausing and genomic compartmentalization. *Nucleic Acids Res.* **42**: 12570–84.

48. **Shang Y**, **Myers M**, **Brown M**. 2002. Formation of the androgen receptor transcription complex. *Mol. Cell* **9**: 601–10.

49. **Johnson KD**, **Christensen HM**, **Zhao B**, **Bresnick EH**. 2001. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol. Cell* **8**: 465–71.

50. **Vieira KF, Levings PP, Hill MA, Crusselle VJ,** et al. 2004. Recruitment of transcription complexes to the β-globin gene locus in Vivo and in Vitro. *J. Biol. Chem.* **279**: 50350–7.

51. **Kim T-K, Hemberg M, Gray JM, Costa AM,** et al. 2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**: 182–7.

52. **De Santa F, Barozzi I, Mietton F, Ghisletti S,** et al. 2010. A Large Fraction of Extragenic RNA Pol II Transcription Sites Overlap Enhancers. *PLoS Biol.* **8**: e1000384.

53. **Koch F, Fenouil R, Gut M, Cauchy P,** et al. 2011. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat. Struct. Mol. Biol.* **18**: 956–63.

54. **Louie MC, Yang HQ, Ma A-H, Xu W,** et al. 2003. Androgen-induced recruitment of RNA polymerase II to a nuclear receptor-p160 coactivator complex. *Proc. Natl. Acad. Sci.* **100**: 2226–30.

55. **Zhao H**, **Dean A**. 2004. An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. *Nucleic Acids Res.* **32**: 4903–14.

56. **Barrow JJ, Li Y, Hossain M, Huang S,** et al. 2014. Dissecting the function of the adult β-globin downstream promoter region using an artificial zinc finger DNA-binding domain. *Nucleic Acids Res.* **42**: 4363–74.

57. **Bulger M**, **Groudine M**. 1999. Looping versus linking: Toward a model for long-distance gene activation. *Genes Dev.* **13**: 2465–77.

58. **Vernimmen D**, **Bickmore WA**. 2015. The Hierarchy of Transcriptional Activation: From Enhancer to Promoter. *Trends Genet.* **31**: 696–708.

59. **Zhu X, Ling J, Zhang L, Pi W,** et al. 2007. A facilitated tracking and transcription mechanism of long-range enhancer function. *Nucleic Acids Res.* **35**: 5532–44.

60. **Wang Q**, **Carroll JS**, **Brown M**. 2005. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol. Cell* **19**: 631–42.

61. **Tuan D**, **Kong S**, **Hu K**. 1992. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc. Natl. Acad. Sci.* **89**: 11219–23.

62. **Szutorisz H, Canzonetta C, Georgiou A, Chow C,** et al. 2005. Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Mol. Cell. Biol.* **25**: 1804–20.

63. **Ho Y**, **Elefant F**, **Liebhaber SA**, **Cooke NE**. 2006. Locus Control Region Transcription Plays an Active Role in Long-Range Gene Activation. *Mol. Cell* **23**: 365–75.

64. **Dillon N, Trimborn T, Strouboulis J, Fraser P,** et al. 1997. The Effect of Distance on Long-Range Chromatin Interactions. *Mol. Cell* **1**: 131–9.

65. **Blackwood EM**, **Kadonaga JT**. 1998. Going the distance: a current view of enhancer action. *Science* **281**: 60–3.

66. **Hatzis P**, **Talianidis I**. 2002. Dynamics of Enhancer-Promoter Communication during Differentiation-Induced Gene Activation. *Mol. Cell* **10**: 1467–77.

67. **Dixon JR, Selvaraj S, Yue F, Kim A,** et al. 2012. Topological domains in

mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376–80.

68. **Nora EP, Lajoie BR, Schulz EG, Giorgetti L,** et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**: 381–5.

69. **Sanborn AL, Rao SSP, Huang S-C, Durand NC,** et al. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci.* **112**: E6456–65.

70. **Rao SSP, Huntley MH, Durand NC, Stamenova EK,** et al. 2014. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* **159**: 1665–80.

71. **Huang J**, **Marco E**, **Pinello L**, **Yuan G-C**. 2015. Predicting chromatin organization using histone marks. *Genome Biol.* **16**: 162.

72. **Melo CA, Drost J, Wijchers PJ, van de Werken H,** et al. 2013. eRNAs Are Required for p53-Dependent Enhancer Activity and Gene Transcription. *Mol. Cell* **49**: 524–35.

73. **Lam MTY, Cho H, Lesch HP, Gosselin D,** et al. 2013. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* **498**: 511–5.

74. **Lai F, Orom UA, Cesaroni M, Beringer M,** et al. 2013. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **494**: 497–501.

75. **Li W, Notani D, Ma Q, Tanasa B,** et al. 2013. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**: 516–20.

76. **Mousavi K, Zare H, Dell'orso S, Grontved L,** et al. 2013. eRNAs Promote Transcription by Establishing Chromatin Accessibility at Defined Genomic Loci. *Mol. Cell* **51**: 606–17.

77. **Schaukowitch K, Joo J-Y, Liu X, Watts JK,** et al. 2014. Enhancer RNA Facilitates NELF Release from Immediate Early Genes. *Mol. Cell* **56**: 29–42.

78. **IIott NE, Heward JA, Roux B, Tsitsiou E,** et al. 2014. Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes. *Nat. Commun.* **5**: 3979.

79. **Hsieh C-L, Fei T, Chen Y, Li T,** et al. 2014. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. *Proc. Natl. Acad. Sci.* **111**: 7319–24.

80. **Sigova AA, Abraham BJ, Ji X, Molinie B,** et al. 2015. Transcription factor trapping by RNA in gene regulatory elements. *Science* **350**: 978–81.

81. **Cloutier SC, Wang S, Ma WK, Al Husini N,** et al. 2016. Regulated Formation of lncRNA-DNA Hybrids Enables Faster Transcriptional Induction and Environmental Adaptation. *Mol. Cell* **61**: 393–404.

82. **Andersson R, Refsing Andersen P, Valen E, Core LJ,** et al. 2014. Nuclear stability and transcriptional directionality separate functionally distinct RNA species. *Nat. Commun.* **5**: 5336.

83. **Lai F**, **Gardini A**, **Zhang A**, **Shiekhattar R**. 2015. Integrator mediates the biogenesis of enhancer RNAs. *Nature* **525**: 399–403.

84. **Paralkar VR, Taborda CC, Huang P, Yao Y,** et al. 2016. Unlinking an lncRNA from Its Associated cis Element. *Mol. Cell* **62**: 104–10.

85. **Kornienko AE**, **Guenzl PM**, **Barlow DP**, **Pauler FM**. 2013. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol.* **11**: 59.

86. **Kolovos P, Knoch TA, Grosveld FG, Cook PR,** et al. 2012. Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin* **5**: 1–8.

87. **Li G, Ruan X, Auerbach RK, Sandhu KS,** et al. 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**: 84–98.

88. **Andersson R**, **Sandelin A**, **Danko CG**. 2015. A unified architecture of transcriptional regulatory elements. *Trends Genet.* **31**: 426–33.

89. **Moorman C, Sun L V., Wang J, de Wit E,** et al. 2006. Hotspots of transcription factor colocalization in the genome of Drosophila melanogaster. *Proc. Natl. Acad. Sci.* **103**: 12027–32.

90. **Xie D, Boyle AP, Wu L, Zhai J,** et al. 2013. Dynamic trans-Acting Factor Colocalization in Human Cells. *Cell* **155**: 713–24.

91. **Chen H, Li H, Liu F, Zheng X,** et al. 2015. An integrative analysis of TFBS-clustered regions reveals new transcriptional regulation models on the accessible chromatin landscape. *Sci. Rep.* **5**: 8465.

92. **Kvon EZ, Stampfel G, Yanez-Cuna JO, Dickson BJ,** et al. 2012. HOT regions function as patterned developmental enhancers and have a distinct cis-regulatory signature. *Genes Dev.* **26**: 908–13.

93. **Smith RP, Taher L, Patwardhan RP, Kim MJ,** et al. 2013. Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. *Nat. Genet.* **45**: 1021–8.

94. **Scholes C**, **DePace AH**, **Sanchez A**. 2016. Integrating regulatory information via combinatorial control of the transcription cycle. *bioRxiv* : 039339.

95. **Levo M**, **Segal E**. 2014. In pursuit of design principles of regulatory sequences. *Nat. Rev. Genet.* **15**: 453–68.

96. **White MA**. 2015. Understanding how cis-regulatory function is encoded in DNA sequence using massively parallel reporter assays and designed sequences. *Genomics* **106**: 165–70.

97. **Inoue F**, **Ahituv N**. 2015. Decoding enhancers using massively parallel reporter assays. *Genomics* **106**: 159–64.

98. **Patwardhan RP, Hiatt JB, Witten DM, Kim MJ,** et al. 2012. Massively parallel functional dissection of mammalian enhancers in vivo. *Nat. Biotechnol.* **30**: 265–70.

99. **Arnold CD, Gerlach D, Stelzer C, Boryn LM,** et al. 2013. Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq. *Science* **339**: 1074–7.

100. **Kwasnieski JC, Mogno I, Myers CA, Corbo JC,** et al. 2012. Complex effects of nucleotide variants in a mammalian cis-regulatory element. *Proc. Natl. Acad. Sci.* **109**: 19498–503.

101. **Sharon E, Kalma Y, Sharp A, Raveh-Sadka T,** et al. 2012. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* **30**: 521–30.

102. **Murtha M, Tokcaer-Keskin Z, Tang Z, Strino F,** et al. 2014. FIREWACh: high-throughput functional detection of transcriptional regulatory modules in mammalian cells. *Nat. Methods* **11**: 559–65.

103. **Dickel DE, Zhu Y, Nord AS, Wylie JN,** et al. 2014. Function-based identification of mammalian enhancers using site-specific integration. *Nat. Methods* **11**: 566–71.

104. **Akhtar W, de Jong J, Pindyurin AV, Pagie L,** et al. 2013. Chromatin Position Effects Assayed by Thousands of Reporters Integrated in Parallel. *Cell* **154**: 914–27.

105. **Melnikov A, Murugan A, Zhang X, Tesileanu T,** et al. 2012. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* **30**: 271–7.

106. **Kheradpour P, Ernst J, Melnikov A, Rogov P,** et al. 2013. Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay. *Genome Res.* **23**: 800–11.

107. **Patwardhan RP, Lee C, Litvin O, Young DL,** et al. 2009. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat. Biotechnol.* **27**: 1173–5.

108. **Gisselbrecht SS, Barrera LA, Porsch M, Aboukhalil A,** et al. 2013. Highly parallel assays of tissue-specific enhancers in whole Drosophila embryos. *Nat. Methods* **10**: 774–80.

109. **Vanhille L, Griffon A, Maqbool MA, Zacarias-Cabeza J,** et al. 2015. High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. *Nat. Commun.* **6**: 6905.

110. **Canver MC, Smith EC, Sher F, Pinello L,** et al. 2015. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**: 192–7.

111. **Ptashne M**, **Gann A**. 1997. Transcriptional activation by recruitment. *Nature* **386**: 569–77.

112. **Stampfel G, Kazmar T, Frank O, Wienerroither S,** et al. 2015. Transcriptional regulators form diverse groups with context-dependent regulatory functions. *Nature* **528**: 1–18.

113. **Keung AJ, Bashor CJ, Kiriakov S, Collins JJ,** et al. 2014. Using Targeted Chromatin Regulators to Engineer Combinatorial and Spatial Transcriptional Regulation. *Cell* **158**: 110–20.

114. **Diao Y, Li B, Meng Z, Jung I,** et al. 2016. A new class of temporarily phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening. *Genome Res.* **26**: 397–405.

115. **Findlay GM, Boyle EA, Hause RJ, Klein JC,** et al. 2014. Saturation editing of genomic regions by multiplex homology-directed repair. *Nature* **513**: 120–3.

116. **Nelles DA, Fang MY, O'Connell MR, Xu JL,** et al. 2016. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* **165**: 488–96.

117. **Chen B, Gilbert LA, Cimini BA, Schnitzbauer J,** et al. 2013. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell* **155**: 1479–91.

118. **Ma H, Tu L-C, Naseri A, Huisman M,** et al. 2016. Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat. Biotechnol.* **34**: 528–30.

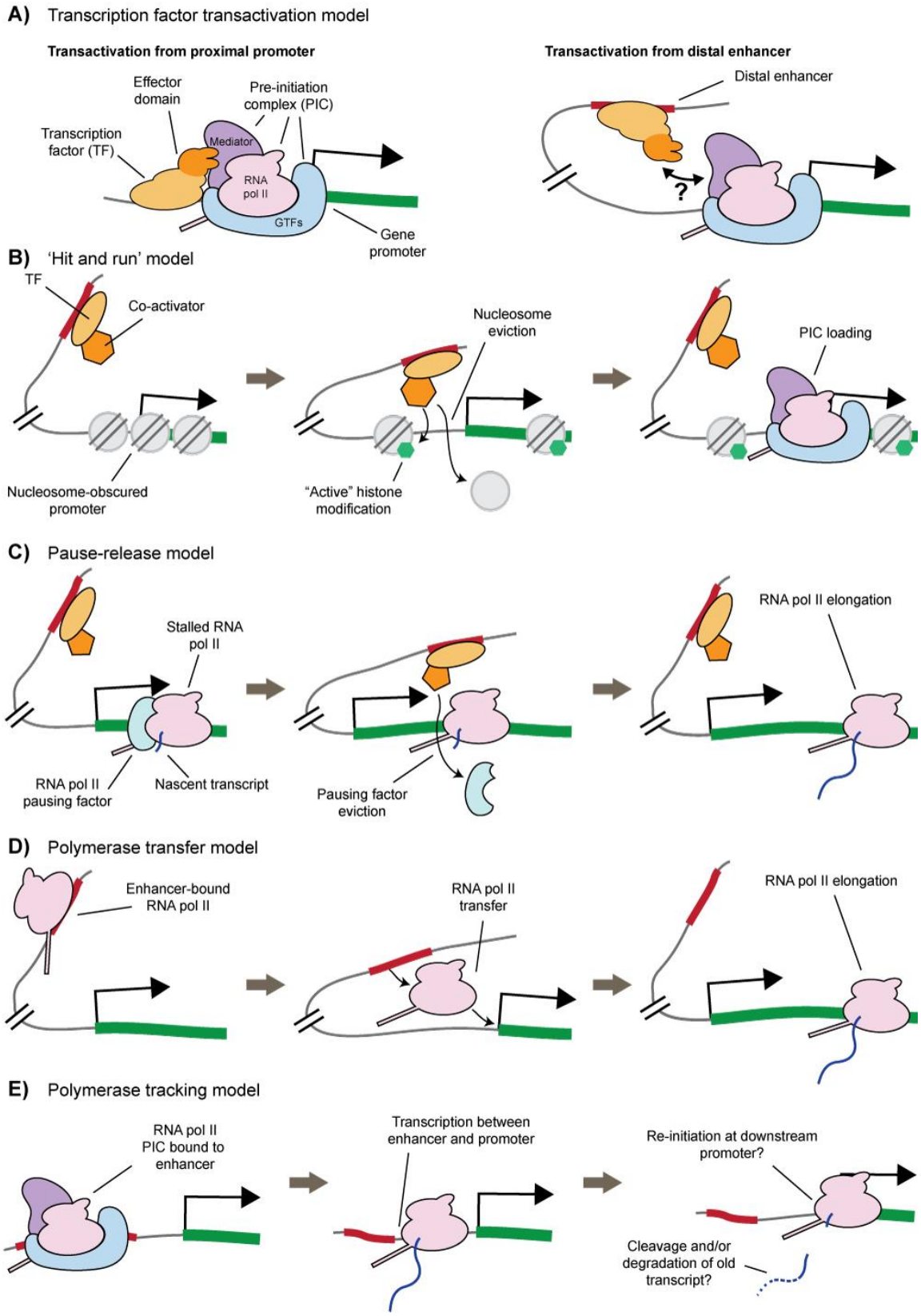
119. **Perez-Pinera P, Kocak DD, Vockley CM, Adler AF,** et al. 2013. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nat. Methods* **10**: 973–6.

120. **Zalatan JG, Lee ME, Almeida R, Gilbert LA,** et al. 2015. Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell* **160**: 339–50.

121. **Shechner DM**, **Hacisuleyman E**, **Younger ST**, **Rinn JL**. 2015. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* **12**: 664–70.

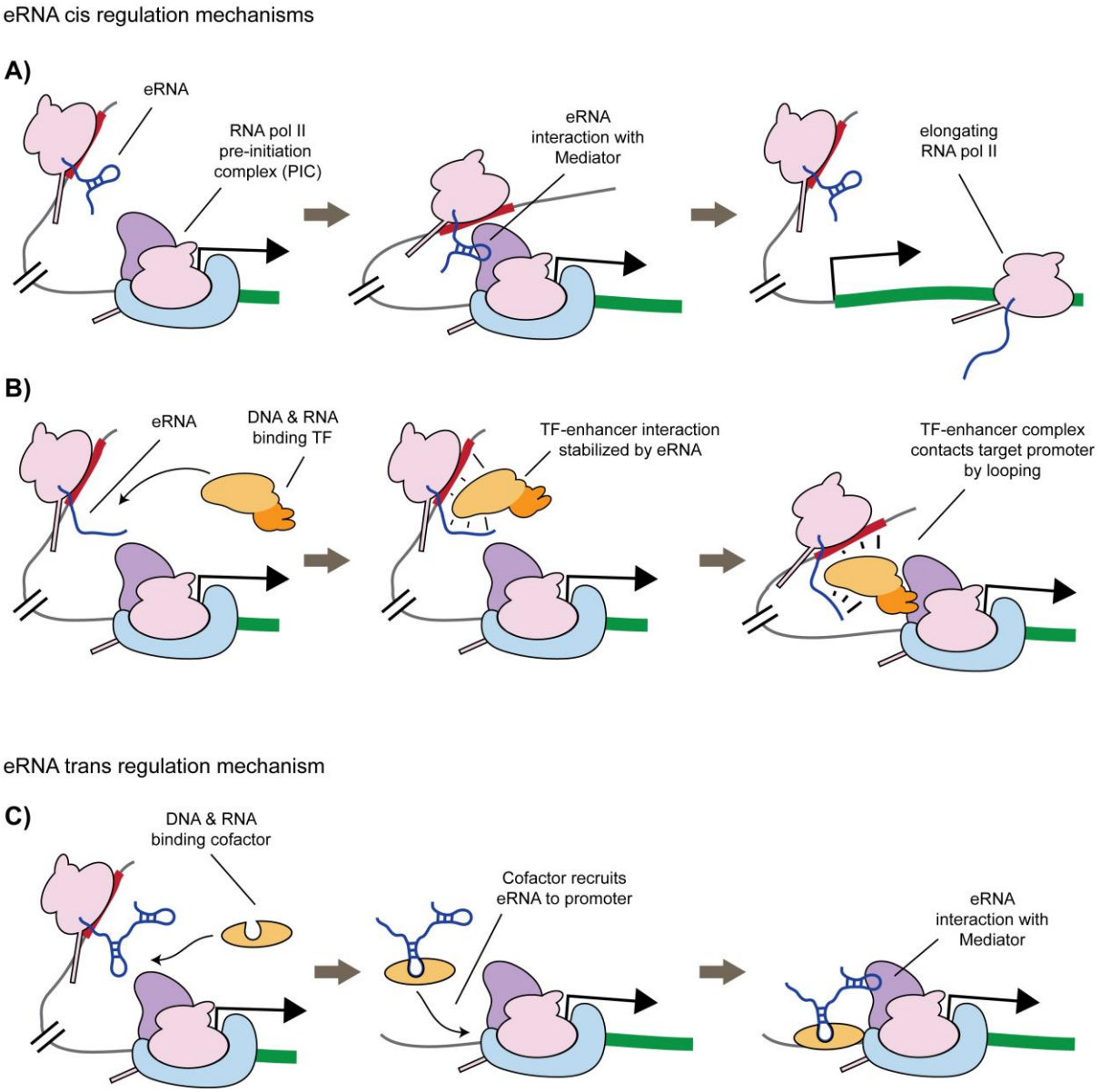
Figures

Figure 1



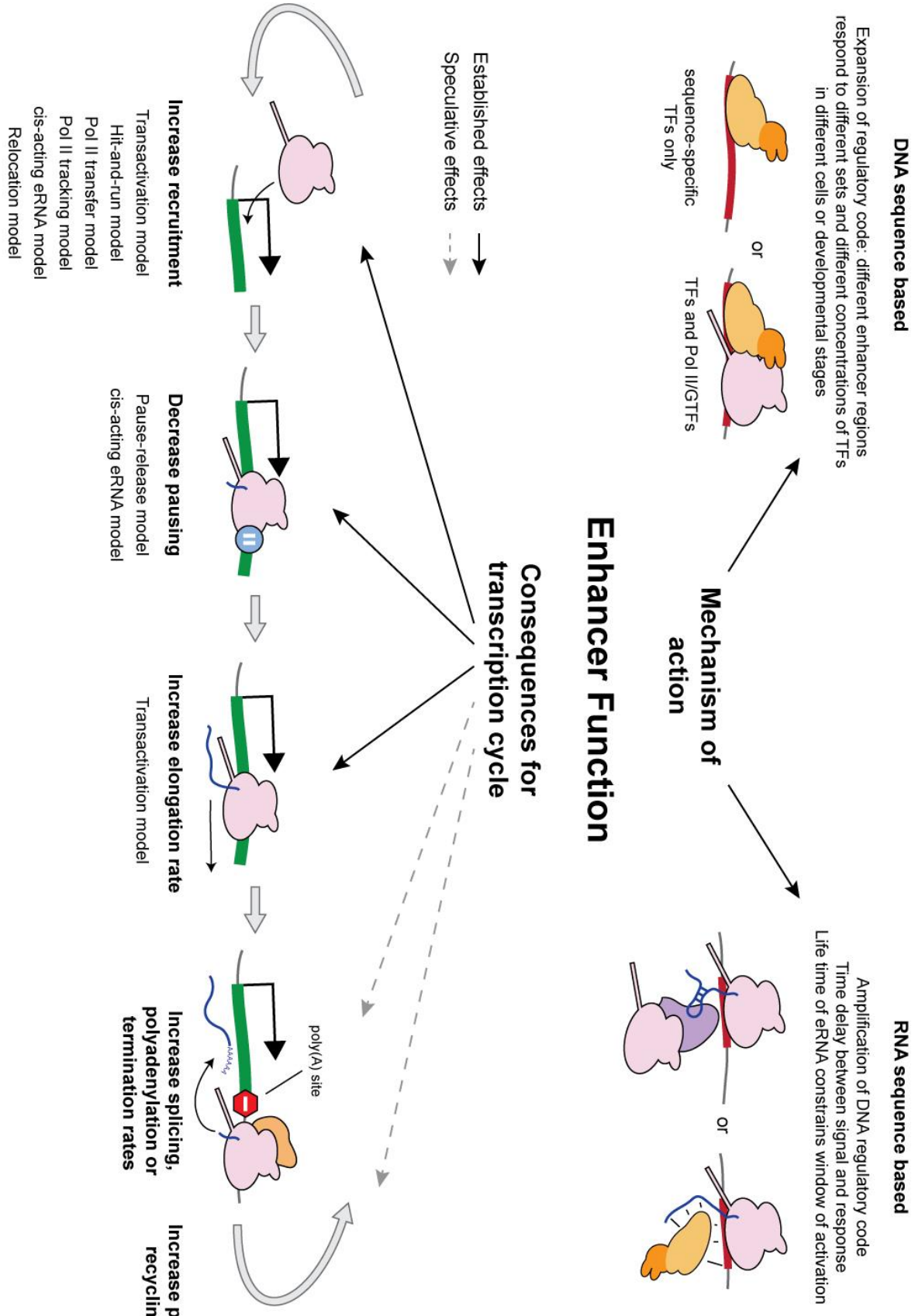
A number of mechanisms have been proposed to explain how enhancers activate transcription of their target genes. **A:** In the transactivation model, transcription factors (TFs) stimulate transcription through interactions with components of the RNA pol II pre-initiation complex (PIC) bound to the gene promoter. Transactivation is well established for TFs that bind directly to proximal promoter sequences (left) but poorly studied for TFs that bind to distal enhancers (right). **B:** In the "hit and run" model, TFs recruit co-activators to enhancer sequences. When the enhancer contacts its target gene, the co-activator is able to generate an open chromatin environment at the promoter. **C:** In the pause-release model, TFs or enhancer-bound co-activators contact target gene promoters and evict pausing factors, allowing RNA pol II to enter into productive transcript elongation. **D:** In the "polymerase transfer" model, RNA pol II bound to the enhancer is physically transferred to a target gene promoter. **E:** RNA pol II loaded at enhancers could locate their target promoters by transcribing through the intervening DNA sequence. The transcripts made during this process must either be degraded co-transcriptionally or cleaved once the promoter is reached. The polymerase could then re-initiate transcription from the gene promoter, or continue transcribing without re-initiation (which would require a distinct mechanism for capping the mRNA transcript).

Figure 2



Transcripts produced from enhancers (eRNAs) might function through cis-acting mechanisms or through trans-acting mechanisms. **A:** Following enhancer-promoter looping, an eRNA could stimulate transcription by interacting directly with components of the RNA pol II pre-initiation complex (PIC). **B:** Alternatively, the RNA transcript might bind to an RNA- and DNA-binding TF, functioning as a "trap" to stabilize the interaction of the TF with its binding site within the enhancer. **C:** eRNAs could act in trans by freely diffusing from the enhancer, followed by recruitment to their target promoter.

Figure 3


Enhancer Function
Mechanism of action
DNA sequence based
Expansion of regulatory code: different enhancer regions respond to different sets and different concentrations of TFs in different cells or developmental stages
sequence-specific TFs only
or
TFs and Pol II/GTFs
RNA sequence based
Amplification of DNA regulatory code
Time delay between signal and response
Life time of eRNA constrains window of activation
or
Consequences for transcription cycle
Established effects
Speculative effects
Increase recruitment
Transactivation model
Hit-and-run model
Pol II transfer model
Pol II tracking model
cis-acting eRNA model
Relocation model
Decrease pausing
Pause-release model
cis-acting eRNA model
Increase elongation rate
Transactivation model
poly(A) site
Increase splicing, polyadenylation or termination rates
Increase pol II recycling

Top: The regulatory information encoded in an enhancer is “read” by binding of TFs or transcriptional co-activators either directly to DNA (left), to transcribed eRNA (right) or to both (far right). Each of these possibilities has distinct implications for the activation of target promoters by the enhancer. **Bottom:** Enhancers can influence gene expression of target genes by affecting different stages of the transcription cycle. Thus far, studies have suggested that enhancers increase the rates of PIC recruitment, transition from paused state to elongation and elongation itself. Theoretically, enhancers could also affect later steps in the transcription cycle (such as polyadenylation and/or termination), although such effects are yet to be observed.