



Transcriptional regulation by promoters with enhancer function

Salvatore Spicuglia, Lan T.M. Dao

► To cite this version:

Salvatore Spicuglia, Lan T.M. Dao. Transcriptional regulation by promoters with enhancer function. Transcription, 2018, 10.1080/21541264.2018.1486150 . hal-01816890

HAL Id: hal-01816890

<https://amu.hal.science/hal-01816890>

Submitted on 15 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Transcriptional regulation by promoters with enhancer function

Lan T.M. Dao¹, Salvatore Spicuglia^{2,3,*}

¹Vinmec Research Institute of Stem cell and Gene technology (VRISG), Hanoi, Vietnam

²Aix-Marseille University, INSERM, TAGC, UMR 1090, Marseille, France.

³Equipe Labéllisée Ligue Contre le Cancer

*Corresponding author

Abstract

Promoters with enhancer activity have been described recently. In this point of view, we will discuss current findings highlighting the commonality of this type of regulatory elements, their genetic and epigenetic characteristics, their potential biological roles in the regulation of gene expression and the underlining molecular mechanisms.

Keywords: Enhancer, Promoter, Epromoter, gene regulation, reporter-assay, CRISPR/Cas9

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Financial disclosure statement

The authors declare that there is no financial disclosure.

Introduction

Gene expression in higher eukaryotes is precisely regulated in time and space through the interplay between promoters and gene-distal *cis*-regulatory regions, known as enhancers ¹. Promoters are located in close proximity to the 5' end of genes and capable of inducing gene expression. Based on the classical definition, enhancers are located distally with respect to the 5' end of genes and activate transcription over large distances and independently of orientation. However, this basic dichotomy of *cis*-regulatory elements has been challenged by broad similarities between genetic and epigenetic properties of promoters and enhancers, including recruitment of the RNA-PolII and general transcription factors, divergent transcription initiation (eRNAs) and context dependent histone modifications ^{2,3}.

Besides the shared architectural characteristics with enhancers, promoters frequently form long-range contacts with other promoters and some promoter elements have been shown to function as enhancers in ectopic enhancer reporter assays ⁴⁻⁶, suggesting that a fraction of promoters might display enhancer functions. Here, we will focus on recent high-throughput functional screens and *in vivo* genetic experiments that have revealed the commonality and physiological functions of these enhancer-like promoters and discuss the potential impact on our current understanding of genome organization and gene expression regulation.

1. Genome-wide identification of promoters with enhancer activity

Survey of the early characterized enhancers using episomal reporter assays revealed that many are located close to, or overlapping with, the promoter region of inducible genes, such as metallothioneins, histones of early cleavage stages, viral immediate-early genes (of some papovaviruses, cytomegaloviruses and retroviruses), heat-shock genes and the antiviral interferon genes ⁶, including the first identified enhancer in the early 80's, which corresponded to the promoter of the Simian Virus 40 (SV40) early gene ⁷.

The development of various powerful techniques that incorporate high-throughput sequencing into reporter assays such as Massively Parallel Reporter Assay (MPRA) and Self-Transcribing Active Regulatory Region sequencing (STARR-seq), have enabled quantitative and straightforward measurements of enhancer activity of thousands of regulatory elements in parallel ^{1,8}. Strikingly, one recurrent observation of these studies is that a number of regions with enhancer activity are actually located proximal or overlapping core promoters ⁹⁻¹⁶.

A screen of the whole *Drosophila* genome using STARR-seq showed that 4.5% of DNA fragments with enhancer activity overlapped with transcription start sites (TSS) ⁹. A subsequent study from the same group compared the specificity of enhancers with the use of different core promoters from either housekeeping or developmentally regulated genes, using a library from

pooled BACs representing ~5 Mb of *D. melanogaster* genomic DNA. Enhancers identified with the housekeeping promoter were preferentially found proximal to or overlapping the TSS of a gene ¹⁰.

Several high-throughput reporter studies in mammals assessing either selected genomic regions (e.g. open chromatin regions or transcription factor binding sites) or human whole genomes have also found a substantial proportion of enhancers overlapping TSS-proximal regions ¹¹⁻¹⁷. In an effort to systematically characterize the type of promoter with enhancer activity, we assessed the enhancer activity of all human core promoters of coding genes by performing CapStarr-seq, a high-throughput reporter assay coupling capture of regions of interest to STARR-seq ¹⁸. For each coding-gene, a region roughly overlapping the core promoter (-200 to +50 bp with respect to the TSS) was captured and the enhancer activity tested in two human cell lines commonly used in the ENCODE project: the cervix adenocarcinoma HeLa-S3 and the erythroleukemia K562 cell lines (**Fig. 1A**) ¹². Remarkably, we found that 2-3% of human promoters displayed enhancer activity in a given cell line ¹², with a significant overlap of enhancer activity between the two cell types. We then denoted this new type of regulatory elements "Epromoters". Consistently, a recent whole genome study in HeLa-S3, using an improved version of the STARR-seq approach, also identified TSS-overlapping enhancers ¹⁴ (**Fig. 2A**; also see below). These reporter-based studies clearly showed that a substantial number of promoters from drosophila to mammals have the intrinsic properties to work as enhancers. This type of regulatory sequences (denoted either Epromoters, TSS-overlapping enhancers or enhancer-like promoters) should be distinguished from promoter-proximal enhancers, which are "classical" enhancers located in proximity, but not overlapping, the core promoter of genes.

2. Distal gene regulation by enhancer-like promoters

High-throughput reporter assays have several intrinsic caveats that might over or underestimate the actual number of promoters with enhancer-like activity ^{1,8}. Besides, the fact that some promoters display enhancer capacity when tested in episomal reporter assays, does not necessarily implies that they actually regulate the expression of distal genes *in vivo*. Therefore, a critical issue is whether some gene promoters are able to function as *bona-fide* enhancers by influencing the activity of other promoters in their endogenous context.

A pioneer study using knockout mouse models reported that intragenic enhancers frequently act as alternative promoters and induce the expression of long non-coding RNA (lncRNA) transcript isoforms ¹⁹. Subsequently, several independent studies using CRISPR genome editing demonstrated that some promoters function as enhancers in their endogenous context. Systematic genomic editing of promoters of lncRNAs co-regulated with neighbored coding

genes demonstrated that the studied lncRNA promoters regulate the expression of the associated neighbor genes²⁰. Further genetic manipulation of these loci showed that regulation of the target genes do not require the lncRNA transcript itself, but instead involves enhancer-like activity of the lncRNA promoters²⁰. Similar finding have been published for other promoters of non-coding transcripts^{19,21}. Using a CRISPR/Cas9-based promoter deletion strategy, we demonstrated that selected promoters of coding genes with enhancer activity, identified in our human STARR-seq reporter assay (i.e. Epromoters), are indeed involved in *cis*-regulation of distal gene expression in their natural context, therefore functioning as *bona fide* enhancers¹² (**Fig. 1B**). These enhancer-like promoters were shown to physically interact with the promoters of the regulated genes, in some cases involving several target genes, implying that in these promoter-promoter interactions, one promoter acts as a distal regulatory element of the other(s). Inversion of one of the model promoters still retained significant enhancer activity, suggesting that, like classical distal enhancers, enhancer-like promoters might display orientation independent enhancer activity.

Two independent studies performing CRISPR/Cas9-based screening of *cis*-regulatory elements found that distal gene promoters control to some extent the expression of other genes^{22,23}. By using a CRISPR/Cas9 epigenetic activation approach, another study showed that two heterologous promoters interacting with the promoter of the T cell inducible gene *CD69* regulate the expression of *CD69* after T cell activation²⁴. Altogether, these studies provided solid evidences that a subset of promoter elements regulate the expression of distal genes *in vivo*, supporting the idea that enhancer-like promoters are involved in complex regulatory networks with interconnected neighbor genes.

3. Characteristics of enhancer-like promoters

In comparison with other promoters, the enhancer-like promoters differ in motif content, transcription factor binding and histone modifications^{11,12,14}. They bind higher levels of p300, a cofactor usually associated with active enhancers²⁵ and display increased ratio of H3K27ac over H3K4me3¹², this ratio correlating with enhancer activity in different cell lines. The enhancer-like promoters are preferentially bound by general inducible transcription factors such as AP1, STAT and ATF/CREB family of transcription factors^{11,12,14}, suggesting a link with housekeeping and stress response functions. Another striking feature of enhancer-like promoters is that they harbor a higher density of distinct motifs and bound transcription factors¹², key properties shared with distal enhancers.

Cis-regulatory elements are commonly associated with transcriptional initiation sites flanking the regulatory sequences². Promoters can be associated with either unidirectional or bidirectional transcription, in the latter the signal intensity being biased towards the sense of

the gene. Enhancers have been shown to generally produce bidirectional unstable transcripts with no particular orientation bias²⁶⁻²⁹ and their relative abundance is positively correlated with enhancer activity²⁷⁻²⁹. A study analyzing transcription initiation at promoters induced during immune challenge in macrophages suggested that the size of the nucleosome-depleted region associated with divergent transcription contribute toward enhancer-like properties³⁰. Indeed, we found in our study that enhancer activity was predominantly detected among bidirectional promoters with unstable divergent transcripts¹², a hallmark of active enhancers²⁶. A recent study developed a transgenic assay in drosophila embryos with dual vectors that simultaneously assesses the ability of the regulatory element to function as an enhancer or a promoter *in vivo*²⁹. This study revealed that gene promoters that were bidirectionally transcribed in the endogenous context displayed both promoter and enhancer activity in the transgenic assay, while unidirectional promoters displayed only promoter activity²⁹. A plausible explanation of these observations will be that divergent transcription at enhancer-like promoters might provide enlarged nucleosome depleted regions serving as a hub for transcription factor binding and establishment of highly active chromatin to further regulate or enhance proximal and distal gene expression.

4. Correlation between promoter and enhancer activities

The fact that enhancer and promoter functions overlap within the same regulatory element rise the question of whether the same regulatory components control either activities or whether these properties can be separated. A study using distinct reporter assays to test for either enhancer or promoter activity of regulatory elements observed a clear positive correlation between both functions¹¹. Furthermore, they used synthetic sequences with tandem repeats of DNA motifs to assess the intrinsic properties of transcription factor binding sites to display promoter or enhancer activities. The study found that distinct DNA motifs were required for either type of activity¹¹. For example, the presence of the AP1 motif resulted in significant enhancer activity, but little promoter activity, while motifs for EGR, CREB, and RFX families of transcription factors generated preferential promoter activity. Thus, it is plausible that within the same regulatory sequence different motifs provide specific enhancer or promoter function.

Another related question is whether promoter and enhancer activities are correlated between them and across different tissues. Enhancer activity of promoter regions assessed by STARR-seq appears to be well correlated with endogenous expression levels in *Drosophila*³¹, but not in human¹². However, some of the human enhancer-like promoters displayed high levels of both promoter and enhancer activity, whereas for others both activities were anti-correlated. Consistently, integrative analysis of epigenomes across human tissues revealed that a given genomic region could have epigenetic features of enhancer or promoter in different tissues,

suggesting that the type of regulatory activity (i.e. enhancer or promoter) might be tissue-specific³². Overall, depending on the locus, enhancer-like promoters might display coordinated enhancer and promoter activities (for example, to control the mRNA expression of clusters of co-regulated genes) or display context-dependent activities.

5. Physiological functions

Enhancer-like promoters appear to be preferentially associated with housekeeping and stress response genes, including interferon response genes^{10,12,14,15}. The study of Zabidi *et al.* showed that enhancers overlapping promoters in *Drosophila* were associated with housekeeping functions¹⁰. Reminiscent of these findings, analysis of random insertion of reporter constructs in *Drosophila* suggested that promoters of housekeeping genes might influence the expression of neighbor loci³³. Barakat *et al.* developed a novel approach combining ChIP enrichment of target genes and STARR-seq in human ES cells¹⁵. They found that the proportion of active enhancers proximal to the TSS was higher for housekeeping genes than for tissue-specific genes¹⁵. By comparing enhancer activity between two mouse cell lines, we also found that TSS-proximal enhancers were less tissue-specific than distal enhancers¹². Overall it appears that enhancer-like promoters might play a role in the regulation of ubiquitous genes, in sharp contrast with the function of distal enhancers that mainly control the expression of tissue specific genes.

The studies from Stark and our laboratories also suggested the involvement of enhancer-like promoters in interferon/stress response^{12,14}. In the context of an inflamed tissue, the inflammatory response mediated by the type I (IFN α /IFN β) and II (IFN γ) interferons (IFNs)-stimulated genes triggers the activation of a complex transcriptional program that is both cell-type- and stimulus-specific and involves the dynamic regulation of hundreds of genes³⁴. In particular, expression of type I interferons can be induced in virtually all cell types upon recognition of viral components, especially nucleic acids. Noteworthy, some cell lines might induce an IFN response following plasmid transfection or have an active interferon response by default, and this in turn could biased the interpretation of high-throughput reporter assays¹⁴.

Enhancer-like promoters in HeLa-S3, but not in K562, were significantly associated with interferon response genes^{12,14}. As pointed out in the above studies, the discrepancy between the two cell lines might come from the fact that HeLa-S3 cells have a basal activation of interferon-response genes. However, when distal interacting genes were taking into account, a significant association with the interferon response was found¹². More generally, stimulation of cell lines with IFN α (including K562) shows a rather specific activation of interferon-

associated enhancer-like promoters, roughly overlapping the ones found in HeLa-S3 cells (unpublished results). Strikingly, we observed that clusters of interferon response genes were generally associated with at least one induced enhancer-like promoter, including several well-known IFN-response clusters such as OAS1-3 or IFIT1-5 family of genes.

To control for potential bias due to the interferon response, Muerdter *et al.* performed whole genome STARR-seq in HeLa-S3 cells treated or not with a combination of two inhibitors of type I interferon response¹⁴. To test whether there is a specific regulation of interferon response genes by enhancer-like promoters we reanalyzed the data from the Muerdter *et al.* study (**Fig. 2A-B**). This study identified 5245 enhancers in untreated HeLa-S3 cells and 9613 enhancers after treatment with the interferon inhibitors¹⁴. In untreated HeLa-S3 cells, active enhancers were significantly over-represented in TSS-proximal regions ($P < 2.2 \times 10^{-15}$ and 1.5×10^{-6} for the locations less than 1kb of the closest TSS of annotated genes, respectively; hypergeometric test; **Fig. 2A**). However, this was not the case after interferon inhibition. While 8% (429) of enhancers were found less than 1kb of the closest TSS in untreated cells, the proportion dropped significantly to 2.6% (247) after interferon inhibition ($P < 2.2 \times 10^{-16}$; Z test for two proportions; **Fig. 2A**). Moreover, functional analyses demonstrated that enhancer-like promoters (enhancers located less than 1kb from the TSS) in untreated HeLa-S3 cells were significantly associated with the interferon response, while distal enhancers were associated with differentiation processes (**Fig. 2B**). Interestingly, after interferon inhibition, the enhancer-like promoters were still enriched for defense response processes. Although the use of interferon inhibitors is recommended to improve the discovery of cell-type specific enhancers¹⁴, the comparison between treated and untreated cells actually provided a strong support to the idea that enhancer-like promoters are functionally linked with interferon response and other mechanisms to fight against pathogens.

Aside the current evidence for the role of enhancer-like promoters in the interferon response, it is likely that this type of regulatory elements could be more generally involved in the rapid response of genes to cellular stress. Indeed, a common characteristic of most of the early-characterized promoters with enhancer activity is that they are associated with inducible genes that have to quickly respond to environmental stress⁶. Therefore, we hypothesize that, within clusters of co-regulated genes, enhancer-like promoters play an essential role in the coordination of rapid gene induction during the inflammatory response, and likely upon other cellular responses to intra- and extra-cellular stress signaling (**Fig. 2C**). This idea fits well with the notion of transcription factories defining discrete membrane-less sub-nuclear compartments containing high concentration of RNA polymerase and transcription factors and where efficient transcription can be triggered^{35,36}. The enhancer-like promoters could either facilitate the assembly or maintenance of the transcription factories by tightening promoter-

promoter interactions or bringing specific transcriptional regulators required for the regulation of the neighbor genes.

6. Perspectives

Provided evidences from high-throughput reporter assays and CRISPR/Cas9-based genetic editing have highlighted the commonality of enhancer-like promoters and suggested that their function might be particularly relevant in the case of rapid and coordinated regulation of gene expression in response to environmental or cellular stimuli. Future work will be required to address the underlying mechanism(s) of action of enhancer-like promoters and to determine whether they define a new type of regulatory elements (i.e., Epromoter) or the "missing link" of a continuum spectrum of *cis*-regulatory elements. Some outstanding questions might involve: What are the specific regulatory components driving promoter *versus* enhancer activity? Can one function be regulated independently of the other? What is the specific contribution of enhancer-like promoters to the regulation of gene expression within the context of transcription factories?

The shared architectural characteristics of enhancers and promoters have also suggested that new genes might originate from enhancer elements ^{2,29,37,38}. As enhancers already have promoter potential (i.e. are transcribed), rapid evolutionary changes in enhancer-flanking DNA (introducing splicing sites or disrupting polyadenylation sites) might give rise to new gene origination. For instances, the majority of hominoid-specific *de novo* protein-coding genes appear to have acquired a regulated transcript structure and expression profile before acquiring coding potential ³⁹. Indeed, analysis of evolutionary conservation revealed that some of these hominoid-specific loci are likely derived from intergenic enhancers in mouse ³⁷. In this context, it will be interesting to examine the dynamic evolution of enhancer-like promoters and assess whether they might represent evolutionary remnants of new genes originated from enhancers.

Finally, these findings have significant implications for the understanding of complex gene regulation in normal development and open the possibility that genetic variants lying within a subset of promoters might influence distal gene expression. In this regards, our study demonstrated that human genetic variation within enhancer-like promoters influence distal gene expression ¹² (**Fig. 1B**). It might be envisioned that disease-associated variants lying within enhancer-like promoters could affect the expression of distal disease-causing genes.

Abbreviations

TSS: transcription start site; IFN: interferon; STARR-seq: Self-Transcribing Active Regulatory Region sequencing; MPRA: Massively Parallel Reporter Assay; ChIP: chromatin

immunoprecipitation; CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats; lncRNA: long non-coding RNA.

Acknowledgements

Work in SS's laboratory is supported by recurrent funding from the Inserm and Aix-Marseille University. The project leading to this publication has received funding from ARC (PJA 20151203149), Plan Cancer (P036496), the Excellence Initiative of Aix-Marseille University - A*Midex, a French "Investissements d'Avenir" program and "Equipe Labellisée Ligue Contre le Cancer" grants. LTMD is supported by Vingroup Joint Stock Company at Vinmec International Hospital, Vietnam.

References

- 1 Chatterjee, S. & Ahituv, N. Gene Regulatory Elements, Major Drivers of Human Disease. *Annual review of genomics and human genetics*, (2017).
- 2 Andersson, R., Sandelin, A. & Danko, C. G. A unified architecture of transcriptional regulatory elements. *Trends in genetics : TIG* **31**, 426-433, (2015).
- 3 Tippens, N. D., Vihervaara, A. & Lis, J. T. Enhancer transcription: what, where, when, and why? *Genes Dev* **32**, 1-3, (2018).
- 4 Li, G. *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84-98, (2012).
- 5 Andersson, R. Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model. *BioEssays : news and reviews in molecular, cellular and developmental biology* **37**, 314-323, (2015).
- 6 Schaffner, W. Enhancers, enhancers - from their discovery to today's universe of transcription enhancers. *Biol Chem* **396**, 311-327, (2015).
- 7 Banerji, J., Rusconi, S. & Schaffner, W. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299-308, (1981).
- 8 Santiago-Algarra, D., Dao, L. T. M., Pradel, L., Espana, A. & Spicuglia, S. Recent advances in high-throughput approaches to dissect enhancer function. *F1000Res* **6**, 939, (2017).
- 9 Arnold, C. D. *et al.* Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074-1077, (2013).
- 10 Zabidi, M. A. *et al.* Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556-559, (2015).
- 11 Nguyen, T. A. *et al.* High-throughput functional comparison of promoter and enhancer activities. *Genome research* **26**, 1023-1033, (2016).
- 12 Dao, L. T. M. *et al.* Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nature genetics* **49**, 1073-1081, (2017).
- 13 Liu, Y. *et al.* Functional assessment of human enhancer activities using whole-genome STARR-sequencing. *Genome Biol* **18**, 219, (2017).
- 14 Muerdter, F. *et al.* Resolving systematic errors in widely used enhancer activity assays in human cells. *Nat Methods* **15**, 141-149, (2018).

- 15 Barakat, T. S. *et al.* Functional dissection of the enhancer repertoire in human embryonic stem cells. *bioRxiv*, (2017).
- 16 Wang, X. *et al.* High-resolution genome-wide functional dissection of transcriptional regulatory regions in human. *bioRxiv*, (2017).
- 17 Ernst, J. *et al.* Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions. *Nat Biotechnol* **34**, 1180-1190, (2016).
- 18 Vanhille, L. *et al.* High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. *Nat Commun* **6**, 6905, (2015).
- 19 Kowalczyk, M. S. *et al.* Intragenic enhancers act as alternative promoters. *Molecular cell* **45**, 447-458, (2012).
- 20 Engreitz, J. M. *et al.* Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*, (2016).
- 21 Paralkar, V. R. *et al.* Unlinking an lncRNA from Its Associated cis Element. *Molecular cell* **62**, 104-110, (2016).
- 22 Diao, Y. *et al.* A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat Methods* **14**, 629-635, (2017).
- 23 Rajagopal, N. *et al.* High-throughput mapping of regulatory DNA. *Nat Biotechnol* **34**, 167-174, (2016).
- 24 Mumbach, M. R. *et al.* Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nature genetics* **49**, 1602-1612, (2017).
- 25 Heintzman, N. D. & Ren, B. Finding distal regulatory elements in the human genome. *Curr Opin Genet Dev* **19**, 541-549, (2009).
- 26 Core, L. J. *et al.* Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nature genetics* **46**, 1311-1320, (2014).
- 27 Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455-461, (2014).
- 28 Henriques, T. *et al.* Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev* **32**, 26-41, (2018).
- 29 Mikhaylichenko, O. *et al.* The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev* **32**, 42-57, (2018).
- 30 Scruggs, B. S. *et al.* Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. *Molecular cell* **58**, 1101-1112, (2015).
- 31 Rennie, S. *et al.* Transcription start site analysis reveals widespread divergent transcription in *D. melanogaster* and core promoter-encoded enhancer activities. *Nucleic acids research*, (2018).
- 32 Leung, D. *et al.* Integrative analysis of haplotype-resolved epigenomes across human tissues. *Nature* **518**, 350-354, (2015).
- 33 Corrales, M. *et al.* Clustering of *Drosophila* housekeeping promoters facilitates their expression. *Genome research* **27**, 1153-1161, (2017).
- 34 Smale, S. T. & Natoli, G. Transcriptional control of inflammatory responses. *Cold Spring Harbor perspectives in biology* **6**, a016261, (2014).
- 35 Feuerborn, A. & Cook, P. R. Why the activity of a gene depends on its neighbors. *Trends in genetics : TIG* **31**, 483-490, (2015).
- 36 Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A Phase Separation Model for Transcriptional Control. *Cell* **169**, 13-23, (2017).

- 37 Wu, X. & Sharp, P. A. Divergent transcription: a driving force for new gene origination? *Cell* **155**, 990-996, (2013).
- 38 Arenas-Mena, C. The origins of developmental gene regulation. *Evol Dev* **19**, 96-107, (2017).
- 39 Xie, C. *et al.* Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs. *PLoS genetics* **8**, e1002942, (2012).
- 40 McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501, (2010).

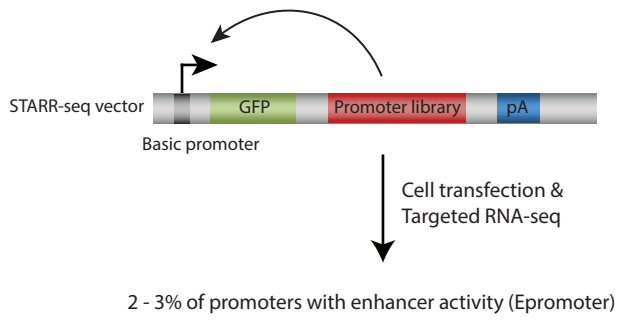
Figure legends:

Figure 1: Functional assessment of promoters with enhancer function. **A)** STARR-seq screening to assess enhancer activity. A human gene promoter library was generated using a capture-based approach and cloned downstream of a core promoter and into the 3'UTR of a GFP reporter gene. After cell transfection, enhancers activate the core promoter and transcribe themselves resulting in reporter transcripts among cellular RNAs. These reporter transcripts are isolated separately by targeted reverse transcription PCR and analyzed by deep sequencing. Among 20,719 coding gene promoters tested in HeLa-S3 and K562 cell lines, we found that 2 to 3% displayed enhancer activity (Epromoters) ¹². **B)** Experimental strategies for the study of Epromoters function in their endogenous context using CRISPR/Cas9 genetic editing. In wild-type cells, an Epromoter (red square) might control the expression of proximal and distal-interacting target genes. When the Epromoter is deleted, the expression of target genes decreased. Enhancer activity is partially retained after inversion of the Epromoter. Allelic replacement of a single nucleotide polymorphism (SNP) overlapping the Epromoter and associated with the expression of a distal gene resulted in impaired expression of the target genes. Dotted lines denote promoter-promoter interactions.

Figure 2: Enhancer-like promoters are preferentially associated with cell defense functions. **A)** Percentage of active enhancer at TSS-proximal regions based on whole genome STARR-seq experiments performed in HeLa-S3 cells in the presence or absence of the TBK1/IKK inhibitor BX-795 and the PKR inhibitor C16 (IFN inhibitors). Data was retrieved from ¹⁴ and reanalyzed using the GREAT tool ⁴⁰. **B)** Enrichment of biological process for TSS-proximal (< 1kb) or distal enhancers identified in the presence or absence of IFN inhibitors. **C)** Model of enhancer-like promoters and gene regulation. Under certain conditions (interferon signaling, stress response or other cell stimulations), chromatin interactions place promoters in close physical proximity (transcription factories), facilitating the recruitment of transcription factors and RNA-PolIII necessary for the transcription of their associated genes. The presence of an enhancer-like promoter (Epromoter) within a cluster of co-regulated genes could facilitate the assembly or maintenance of the transcription factory by tightening the promoter-promoter interactions or bringing specific transcriptional regulators required for the regulation of the neighbor gene.

Figure 1

A



B

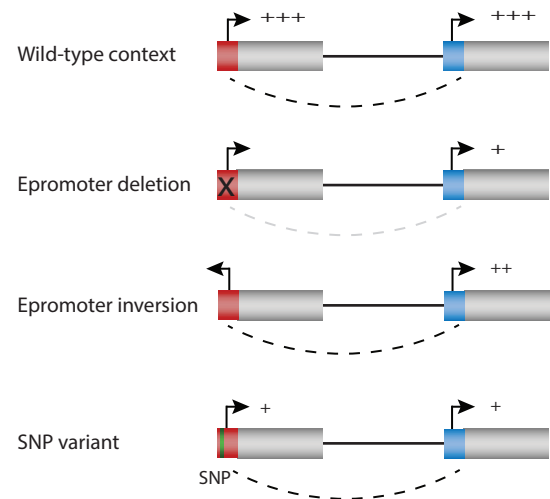


Figure 2

