



**HAL**  
open science

## Recent advances in high-throughput approaches to dissect enhancer function

David Santiago-Algarra, Lan T. M. Dao, Lydie Pradel, Alexandre España, Salvatore Spicuglia

► **To cite this version:**

David Santiago-Algarra, Lan T. M. Dao, Lydie Pradel, Alexandre España, Salvatore Spicuglia. Recent advances in high-throughput approaches to dissect enhancer function. *F1000Research*, 2017, 6, pp.939. 10.12688/f1000research.11581.1 . hal-01612096

**HAL Id: hal-01612096**

**<https://amu.hal.science/hal-01612096v1>**

Submitted on 28 May 2024

**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



## REVIEW

# Recent advances in high-throughput approaches to dissect enhancer function [version 1; peer review: 3 approved]

David Santiago-Algarra, Lan T.M. Dao, Lydie Pradel, Alexandre España ,  
Salvatore Spicuglia 

Aix-Marseille University, TAGC, Marseille, France

**V1** **First published:** 19 Jun 2017, 6(F1000 Faculty Rev):939  
<https://doi.org/10.12688/f1000research.11581.1>  
**Latest published:** 19 Jun 2017, 6(F1000 Faculty Rev):939  
<https://doi.org/10.12688/f1000research.11581.1>

## Abstract

The regulation of gene transcription in higher eukaryotes is accomplished through the involvement of transcription start site (TSS)-proximal (promoters) and -distal (enhancers) regulatory elements. It is now well acknowledged that enhancer elements play an essential role during development and cell differentiation, while genetic alterations in these elements are a major cause of human disease. Many strategies have been developed to identify and characterize enhancers. Here, we discuss recent advances in high-throughput approaches to assess enhancer activity, from the well-established massively parallel reporter assays to the recent clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9-based technologies. We highlight how these approaches contribute toward a better understanding of enhancer function, eventually leading to the discovery of new types of regulatory sequences, and how the alteration of enhancers can affect transcriptional regulation.

## Keywords

gene transcription, enhancer function, MPRA, STARR-seq, CRISPR

## Open Peer Review

**Approval Status** 

	1	2	3
<b>version 1</b> 19 Jun 2017			

**Faculty Reviews** are review articles written by the prestigious Members of **Faculty Opinions**. The articles are commissioned and peer reviewed before publication to ensure that the final, published version is comprehensive and accessible. The reviewers who approved the final version are listed with their names and affiliations.

1. **Alvaro Rada-Iglesias**, University of Cologne, Cologne, Germany
2. **Douglas Vernimmen**, University of Edinburgh, Edinburgh, UK
3. **Bart Deplancke**, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Any comments on the article can be found at the end of the article.

**Corresponding author:** Salvatore Spicuglia ([salvatore.spicuglia@inserm.fr](mailto:salvatore.spicuglia@inserm.fr))

**Competing interests:** The authors declare that they have no competing interests.

**Grant information:** Work in Salvatore Spicuglia's laboratory was supported by recurrent funding from the Inserm and Aix-Marseille University and by the ARC (PJA 20151203149) and the Plan Cancer (P036496) grants. David Santiago-Algarra and Lan T.M. Dao were supported by the CONACYT-Mexico and Vietnam International Education Development, respectively.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2017 Santiago-Algarra D *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Santiago-Algarra D, Dao LTM, Pradel L *et al.* **Recent advances in high-throughput approaches to dissect enhancer function [version 1; peer review: 3 approved]** F1000Research 2017, 6(F1000 Faculty Rev):939 <https://doi.org/10.12688/f1000research.11581.1>

**First published:** 19 Jun 2017, 6(F1000 Faculty Rev):939 <https://doi.org/10.12688/f1000research.11581.1>

## Introduction

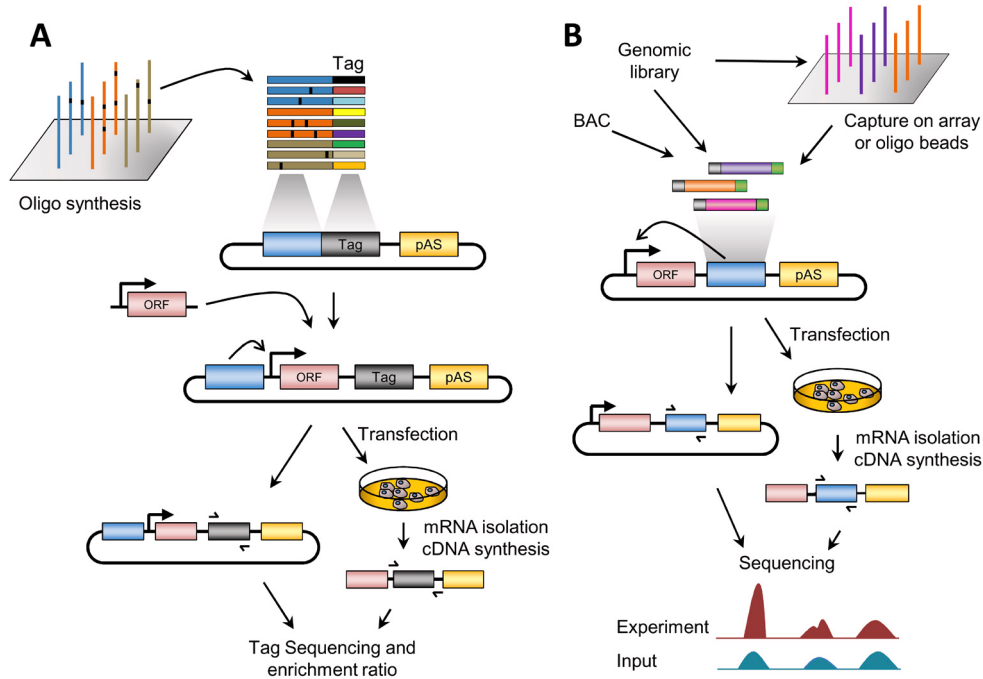
Gene expression is precisely regulated by a combination of promoters and gene-distal regulatory regions, known as enhancers<sup>1,2</sup>. With the increasing awareness of the important role of enhancers in normal development as well as in disease, there is strong scientific interest in identifying and characterizing these elements. This is a challenging task because an enhancer does not have to be located directly adjacent to the gene it regulates. Putative enhancers can be identified across entire genomes based on open chromatin regions (e.g. based on DNase I-seq or ATAC-seq) or chromatin signatures (H3K4me1, H3K27ac), which map the potentially active enhancers<sup>3</sup>. Although useful, these approaches do not provide direct proof of enhancer function, nor do they allow insights into the discrete sequences required for enhancer activity. Therefore, it is crucial to test whether genomic regions actually function as *bona fide* enhancers in living cells or tissues.

In recent years, various powerful techniques that incorporate high-throughput sequencing into reporter assays have enabled quantitative and straightforward measurements of enhancer activity

of thousands of regulatory elements. More recently, the advent of clustered regularly interspaced short palindromic repeats (CRISPR)-related approaches allows massively assessing the relevance of enhancer function in the endogenous context. This review summarizes the assays developed for functional genome-wide testing of enhancer activity and their limitations as well as the main findings that have been gathered using these techniques.

## Principle of high-throughput reporter assays

Episomal reporter assays have been widely used to characterize putative regulatory regions. Several high-throughput strategies have been developed, permitting the simultaneous analysis of hundreds of thousands of reporter plasmids at once. These have been the focus of several comprehensive reviews (e.g. 4–6). These methods can be either qualitative (usually based on cell sorting) or quantitative (based on RNA-seq) and designed to test enhancer or promoter activity. Here, we will focus on recent quantitative methods aiming to characterize enhancers. In particular, two approaches have been widely used in recent years (Figure 1; Table 1): massively parallel reporter assay (MPRA) and self-transcribing active regulatory region sequencing (STARR-seq).



**Figure 1. Principle of high-throughput assays for enhancer activity.** (A) Overview of massively parallel reporter assay (MPRA). The test sequences (wild-type, variants, etc.) are generally synthesized *in silico* by massive oligonucleotide synthesis with unique barcode tags and cloned into the plasmid backbone. Tags can be synthesized along with the test sequences or added after synthesis by polymerase chain reaction (PCR) amplification. A basal promoter and a reporter open reading frame (ORF) are inserted between the tested element and tag sequences. The reporter library is then transfected into cultured cells. Subsequently, mRNA is isolated and cDNA synthesized. The tags are sequenced before (plasmid library pool, for normalization) and after the transfection. The difference in the enrichment of each barcode is proportional to the enhancer activity of the test sequence. In the case of post-synthesis addition of barcodes, an additional sequencing step is required at the first cloning step. (B) Overview of self-transcribing active regulatory region sequencing (STARR-Seq). A genomic or bacterial artificial chromosome (BAC) library is cloned in the reporter plasmid, downstream of the ORF and upstream of the polyadenylation site (pAS). Alternatively, the regions of interest might be enriched by a capture approach. The reporter library is transfected into cultured cells. Subsequently, mRNA is isolated and cDNA synthesized. The cloned regions are sequenced from the plasmid library pool (input) and the cDNA. Differences in the enrichment with respect to the input are proportional to the enhancer activity. In both panels, the effect of the enhancer on the basal promoter is indicated by an arrow.

**Table 1. Examples of high-throughput functional assays of mammalian enhancers.**

Original technique	Specific Name	DNA origin	Application	Specific features	No. of regions <sup>1</sup>	Size (bp)	Cell lines	Promoter	Species	Ref.
MPRA		Synthetic	Characterize putative enhancers and their active region	Centered on TF sites	2,104	145	K562, HepG2	SV40	Human	12
MPRA	CRE-seq	Synthetic	Characterize genomic regions with predicted function as enhancers	Mutation of TF binding sites	2,100	130	K562	Hsp68	Human	9
FIREWACH		Genomic	Identification of enhancers from restriction enzyme accessible regions	Isolation of GFP-positive regions	84,240	154	ESC	Minimal FpG5	Mouse	25
STARR-seq	CapSTARR-seq	Genomic	Identification of enhancers from DHS	Capture of DHS	7,542	330–430	P5424, 3T3	SCP1	Mouse	22
STARR-seq		Genomic	Characterize genetic variants in regulatory elements from 95 subjects	Capture of DHS	104	402	HepG2	SCP1	Human	23
MPRA	CRE-seq	Genomic	Identification of regulatory elements	Capture of DHS	4,000	464	Retina	Minimal Rho	Mouse	24
MPRA	Sharp-MPRA	Synthetic	Characterize nucleotides as activators or repressors in regulatory elements	Tiled oligonucleotide synthesis (30 or 5 bp resolution)	15,720	145	HepG2, K562	Minimal TATA, SV40	Human	13
MPRA		Synthetic	Identification of regulatory variants associated with eQTLs	Centered on eQTL variants	3,642	150	Lymphoblastoid, HepG2	Minimal TATA	Human	14
MPRA		Synthetic	Identification of regulatory variants associated with red blood cell traits	Three sliding windows with respect to the GWAS variants	2,756	145	K562, K562+GATA1	Minimal TATA	Human	15
STARR-seq	CapSTARR-seq	Genomic	Identification of promoter regions with enhancer activity	Capture of -200 to +50 bp with respect to the TSS of coding genes	20,719	250	K562, HeLa	SCP1	Human	33
MPRA	LentiMPRA	Synthetic	Identification of enhancers by chromosomal integration	Centered on several ChIP-seq peak signals	2,236	171	HepG2	pGL4.23 promoter	Human	26
CRISPR-Cas9		Synthetic	Identify endogenous enhancers bound by p53 and ER $\alpha$ TFs	Target p53 binding site in putative enhancers	685	N/A	BJ-RASg12v	N/A	Human	44

<sup>1</sup>Number of targeted DNA sequences, not necessarily the number of unique fragments that are tested

Bp, base pair; capSTARR-seq, capture-based self-transcribing active regulatory region sequencing; CHIP-seq, chromatin immunoprecipitation sequencing; CRE, cis-regulatory elements; CRISPR, clustered regularly interspaced short palindromic repeats; DHS, DNase I hypersensitive sites; eQTL, expression quantitative trait loci; ER $\alpha$ , estrogen receptor alpha; ESC, embryonic stem cell; FIREWACH, functional identification of regulatory elements within accessible chromatin; GFP, green fluorescent protein; GWAS, genome-wide association study; Hsp68, heat shock promoter 68; lentiMPRA, lentiviral massively parallel reporter assay; MPRA, massively parallel reporter assay; N/A, not applicable; Sharp, systematic high-resolution activation and repression profiling with reporter-tiling; STARR-seq, self-transcribing active regulatory region sequencing; SCP1, super core promoter 1; SV40, simian virus 40; TF, transcription factor; TSS, transcription start site.

The MPRA method consists of the generation of a library of reporter constructs based on microarray synthesis of DNA sequences (generally, tested sequences are cloned upstream of a basal promoter) and unique sequence tags or barcodes (placed in the 3' UTR of the reporter gene). To increase the sensitivity and reproducibility, several barcodes could be added to any given sequence. The reporter library is then transfected into cell lines of interest and RNA sequencing of the barcodes is performed, thus providing a quantitative readout of the regulatory activity of the tested regions (Figure 1A).

MPRAs have been used to investigate a number of biological questions. Initially, MPRA was designed to dissect the functional components of previously identified enhancers at single-nucleotide resolution<sup>7,8</sup>. Subsequently, a similar approach (also named CRE-seq) was used to functionally test ~2,000 genomic segments predicted by ENCODE to be enhancers, weak enhancers, or repressed elements<sup>9</sup> as well as test synthetic enhancers to model grammatical rules of regulatory sequences<sup>10,11</sup>. MPRA can be used to systematically assess the relevance of predicted regulatory motifs within enhancers. Kheradpour *et al.* tested ~2,000 predicted enhancers along with engineered enhancer variants containing targeted motif disruptions for key transcription factors (TFs)<sup>12</sup>. In a follow-up study, Kellis' lab developed a high-resolution MPRA approach (also named Sharpr-MPRA) that allowed genome-scale mapping of activating and repressive nucleotides in regulatory regions<sup>13</sup>. Here, by synthesizing dense tiling of overlapping MPRA constructs, they managed to infer the regulatory effects of functional regulatory nucleotides with either activating or repressive properties<sup>11,13</sup>. Finally, MPRA can be used to test the impact of single nucleotide polymorphisms (SNPs) in order to identify functional regulatory variants linked to human traits or diseases. Two recent studies from the Broad Institute provide proof-of-concept for such approaches. Tewhey *et al.* used an improved version of the MPRA to analyze thousands of human expression quantitative trait loci (eQTL) to identify alleles that impact gene expression in lymphoblastoid cell lines<sup>14</sup>. Ulirsch *et al.* used MPRA to test 2,756 variants linked to 75 genome-wide association studies (GWAS) loci involved in red blood cell traits<sup>15</sup>. In both cases, CRISPR-mediated genetic engineering confirmed the relevance of the MPRA findings. Interestingly, some of the identified regulatory variants did not lie within known motifs, suggesting that they can influence DNA structure instead of binding of TFs<sup>15</sup>.

A new innovative method named STARR-seq was introduced by Alexander Stark and colleagues<sup>16</sup>. STARR-seq is an MPRA (reviewed in 17) aimed to identify and quantify transcriptional enhancers directly based on their activity in entire genomes (Figure 1B). In brief, a bulk of DNA fragments from arbitrary sources is cloned downstream into the 3' UTR of a GFP reporter gene. Once in cellular context, active enhancers will activate the upstream promoter and transcribe themselves, resulting in reporter transcripts among cellular RNAs. Thus, each reporter transcript contains the reporter gene and the "barcode" of itself. These reporter transcripts can be isolated separately by targeted PCR and eventually detected by high-throughput sequencing. In this way, the activity of millions of putative enhancers can be measured

simultaneously without being affected by the location of the candidate sequences and their orientation. The main advantage over the classical MPRA is that the tested sequence itself is used as a "barcode", substantially simplifying the whole procedure to quantify enhancer activity. Stark's lab used the STARR-seq approaches to ask several basic mechanistic questions of enhancer biology in *Drosophila*, including (i) identification and characterization of cell-type-specific<sup>16,18</sup> and hormone-responsive enhancers<sup>19</sup>, (ii) the impact of *cis*-regulatory sequence variation on enhancer activity and evolution<sup>20</sup>, and (iii) dissecting the basis of enhancer core-promoter specificity<sup>21</sup>.

STARR-seq has been applied to human cells by utilizing selected bacterial artificial chromosomes (BACs)<sup>16</sup>; however, with the complexity and size of mammalian genomes, this technique is not easily implemented, making the formulation of representative libraries a challenge and a very high sequencing depth a necessity. To avoid this issue, we developed a capture-based approach (named CapSTARR-seq) to assess a subset of mouse DNase I hypersensitive sites (DHSs) found in developing thymocytes<sup>22</sup>. Here, the regions of interest are captured using custom-designed microarrays and cloned into the STARR-seq vector, thus providing cost-effective and accurate quantification of enhancer activity in mammals. Similar approaches have been published by other labs, including capture of natural genomic variants<sup>23</sup> and test of DHSs from the central nervous system using a capture approach with oligo-baits<sup>24</sup>. Alternatively, it could be possible to directly clone open chromatin regions, as described in the functional identification of regulatory elements within accessible chromatin (FIREWACH) method<sup>25</sup>.

### Potential caveats of high-throughput reporter assays

The DNA sources used for testing are a potential issue of high-throughput reporter assays. Most MPRA approaches have used massive oligonucleotide synthesis (Figure 1A), which allows the precise definition of tested regions as well as custom modifications of underlying sequences. However, there are currently two limitations to this approach. On the one hand, the size of the tested fragment is limited to ~200 bp (including the adaptors), which might prevent testing the full regulatory regions. On the other hand, there is a limitation in the number of oligonucleotides that can be synthesized (currently <100,000). These constraints are expected to be overcome in the near future with the improvement of oligonucleotide synthesis technologies.

In the STARR-seq approach, the DNA fragments are cloned within the transcribed region (Figure 1B), which is very convenient because their sequences provide direct information about enhancer activity. However, it also introduces a source of potential artifacts, as some sequences might influence transcript stability instead of enhancing transcription. This potential bias could be avoided by comparing the results of tested regions on both orientations, allowing one to filter out the effects of strand-specific transcript-stabilizing effects.

A general concern about the episomal reporter assays is that they may not accurately reflect the function of enhancer elements in their endogenous context. To partially circumvent this caveat,

chromatinized adeno-associated<sup>11,24</sup> and lentiviral MPRA<sup>25,26</sup> have also been performed. These methods capacitate reporter assays within cells or tissues that are difficult to transfect. Certainly, an equally valid argument is that episomal reporter assays allow the unbiased study of enhancer function independently of any “perturbing” chromatin or genomic context. Interestingly, a recent study performed a systematic comparison of chromosomal versus episomal enhancer activity using integrative and non-integrative versions of a lentiviral-based reporter assay<sup>26</sup>. Although the chromosomally based reporter assay was more predictable by epigenomic and sequence-based models, both reporter assays remained relatively well correlated. Another alternative approach is the introduction of reporter genes throughout the genome using transposition systems (e.g. 27–29). Although these approaches do not directly assess enhancer activity, they allow one to infer the regulatory context of endogenous loci.

### Some relevant findings

One interest of high-throughput enhancer assays is the possibility to explore enhancer function without preconceived notions, thus potentially leading to new unforeseen findings. A common observation of several studies is that many predicted enhancer regions did not show reporter activity<sup>9,12,13,22,26</sup>. For example, only 26% of predicted enhancers based on chromatin signatures in K562 cells displayed enhancer activity in the reporter assays performed in the same line<sup>9</sup>, suggesting that, in addition to histone modifications, additional sequence specificity, such as TF-binding sites, are essential determinants of *cis*-regulatory activity. Indeed, the concentration of TF-binding sites or motifs is highly predictive of strong enhancer activity<sup>13,22,30</sup>. Alternatively, this could also indicate that not all of the required sequences are present in the tested regions or that endogenous promoter contexts are essential to the enhancer activity. Finally, it is also plausible that some open chromatin regions, while contributing to transcriptional regulation, have enhancer-independent functions<sup>31</sup> or lack classical enhancer functions<sup>32</sup>.

Perhaps the most intriguing finding of functional enhancer assays comes from the observation that many core promoter regions display enhancer activity<sup>11,16,21,33</sup>. The original definition of enhancers implies the ability to activate gene expression at a distance, while promoters entail the capability to induce local gene expression. However, this basic dichotomy of *cis*-regulatory elements has been challenged by broad similarities between promoters and enhancers, such as DNA sequence features, chromatin marks, Pol II recruitment, and bidirectional transcription<sup>34</sup>. For instance, H3K4me3, a histone modification generally found at promoter regions, has been also associated with active enhancers<sup>35–37</sup>. Assessment of enhancer activity by Cap-STARR-seq showed that strong transcription start site (TSS)-distal enhancers are indeed associated with H3K4me3 enrichment at the endogenous loci<sup>22</sup>. Several studies have also suggested that some promoters might play enhancer functions<sup>34</sup>. The extent of this type of promoter and whether it actually functions to regulate the expression of distal genes have remained elusive. Now, several independent studies based on massive reporter assays reported widespread enhancer activity from TSS-proximal regions. By

applying STARR-seq, Zabidi *et al.* screened the whole fly genome with the use of different core promoters obtained from either ubiquitously expressed housekeeping genes or developmentally regulated and cell-type-specific genes<sup>21</sup>. They found that promoters of housekeeping genes were mainly regulated by promoter-proximal enhancers, while promoters of developmental and cell-type-specific genes required distally located enhancers. Ernst *et al.* found that active enhancers were enriched in DNase I sites overlapping TSS in human cell lines<sup>13</sup>. Nguyen *et al.* performed a functional comparison of a subset of promoters and enhancers in mouse neurons using an integrative MPRA approach<sup>11</sup>. Interestingly, gene promoters generated similar enhancer activity as compared to distal regulatory regions. In a recent study, we found that 2–3% of all human core promoters displayed enhancer activity in a given cell line<sup>33</sup>. Compared to classical promoters and distal enhancers, these TSS-overlapping enhancers displayed distinct genomic and epigenomic features and were associated with housekeeping and stress response genes. CRISPR genomic deletions demonstrated that several core promoters with enhancer activity in the reporter assay are indeed involved in *cis*-regulation of distal gene expression in their natural context, therefore functioning as *bona fide* enhancers. Furthermore, human genetic variation within this type of promoter was associated with a strong effect on distal gene expression. Concomitantly, another study, using comprehensive genetic manipulation of promoter regions, reported frequent distal *cis*-regulation by loci associated with promoters of lncRNAs and, to a lesser extent, coding genes<sup>32</sup>. Finally, two recent studies performing screens of *cis*-regulatory elements by CRISPR/Cas9-based approaches (see below) have found that the expression of some genes is controlled by distal gene promoters<sup>38,39</sup>. Overall, these findings open up the intriguing possibility that developmental traits or disease-associated variants lying within a subset of promoters might directly impact on distal gene expression.

### CRISPR-based approaches to assess enhancer function

As discussed above, a potential limitation of reporter assays is that candidate enhancers are studied outside their endogenous genomic context, which is likely required for their *in vivo* function. The advent of CRISPR-based technologies now allows the circumvention of this caveat. Several studies have performed systematic dissection of individual enhancers using either TALEN<sup>40</sup> or CRISPR-mediated mutagenesis<sup>41</sup>. In these studies, a tiling single guide RNA (sgRNA), or TALEN, library covering selected enhancers was designed to perform *in situ* saturating mutagenesis screens, pinpointing sequences with either positive or negative impact on enhancer function. The CRISPR approach was subsequently extended to assess enhancer function within large genomic regions surrounding key loci<sup>38,39,42,43</sup> or to screen for enhancer elements involved in specific gene regulation pathways<sup>44</sup> (Table 1).

CRISPR-mediated mutagenesis is limited by the fact that a high density of sgRNAs is required to saturate all possible regulatory elements and specific mutations are difficult to implement genome-wide. In some cases, there might also be a bias with respect to

the regions that can be targeted by the designed sgRNAs or limitations owing to transfection efficiency in particular cell types. Alternatively, nuclease-deactivated Cas9 (dCas9) can be fused to activator or repressor domains to precisely modify gene expression from promoters and distal regulatory elements<sup>45</sup>. Based on this property, repressor and activator domains fused with dCas9 combined with a pool of sgRNAs have been used for comprehensive CRISPR-inactivation (CRISPRi) and CRISPR-activation screens targeting DHSs of a gene of interest<sup>46</sup> or an entire locus<sup>47</sup>. In the former study, a reporter gene introduced at the place of the target genes was used to monitor enhancer activity. In the latter study, the screening criteria were based on the growth advantage or disadvantage provided by the change in expression of the enhancer-associated gene, thus providing proof-of-concept for screening of functional regulatory regions genome-wide. A current limitation of these approaches is that the screening strategy might be based on phenotypic features (such as cell growth fitness, developmental markers, etc.) instead of directly assessing the expression levels of regulated genes. To overcome this limitation, a new powerful method combined CRISPRi and single-cell RNA-seq<sup>48</sup>, enabling high-throughput interrogation of enhancers at single-cell resolution and directly linking enhancer function, and their combinations, with its target gene(s). Although these approaches have been used so far to scan restricted genomic areas, they will likely be implemented in true genome-wide screens of regulatory elements in the coming future.

### Concluding remarks

The implementation of high-throughput reporter assays and CRISPR-based screens allows the experimental validation of enhancer activity in different cell types and cellular contexts (Table 1). These assays are now robust and sensitive enough to be widely used as part of the toolkit for researchers interested in gene regulation. These approaches also led to unpredicted

discoveries, such as the role of core promoters as enhancer-like regulators. One main limitation of these approaches remains the fact that they do not provide direct information towards uncovering enhancer target genes. Therefore, the combination of enhancer assays with recently developed 3C-related methodologies, such as 4C-seq, Hi-C or Capture Hi-C<sup>49</sup>, should greatly facilitate the assignment of discovered enhancers to their putative target genes. Finally, with the expected decrease in the cost of sequencing and oligo synthesis, it will be possible to systematically test the impact of regulatory variants in different diseases and developmental contexts.

### Abbreviations

CRISPR, clustered regularly interspaced short palindromic repeats; DHS, DNase I hypersensitive site; MPRA, massively parallel reporter assay; sgRNA, single guide RNA; STARR, self-transcribing active regulatory region; TF, transcription factor; TSS, transcription start site.

### Competing interests

The authors declare that they have no competing interests.

### Grant information

Work in Salvatore Spicuglia's laboratory was supported by recurrent funding from the Inserm and Aix-Marseille University and by the ARC (PJA 20151203149) and the *Plan Cancer* (P036496) grants. David Santiago-Algarra and Lan T.M. Dao were supported by the CONACYT-Mexico and Vietnam International Education Development, respectively.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

### References

- Kim TK, Shiekhattar R: **Architectural and Functional Commonalities between Enhancers and Promoters.** *Cell.* 2015; **162**(5): 948–59.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Vernimmen D, Bickmore WA: **The Hierarchy of Transcriptional Activation: From Enhancer to Promoter.** *Trends Genet.* 2015; **31**(12): 696–708.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Chatterjee S, Ahituv N: **Gene Regulatory Elements, Major Drivers of Human Disease.** *Annu Rev Genomics Hum Genet.* 2017; **18**.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Dailey L: **High throughput technologies for the functional discovery of mammalian enhancers: new approaches for understanding transcriptional regulatory network dynamics.** *Genomics.* 2015; **106**(3): 151–8.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Inoue F, Ahituv N: **Decoding enhancers using massively parallel reporter assays.** *Genomics.* 2015; **106**(3): 159–64.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- White MA: **Understanding how cis-regulatory function is encoded in DNA sequence using massively parallel reporter assays and designed sequences.** *Genomics.* 2015; **106**(3): 165–70.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Patwardhan RP, Hiatt JB, Witten DM, *et al.*: **Massively parallel functional dissection of mammalian enhancers *in vivo*.** *Nat Biotechnol.* 2012; **30**(3): 265–70.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Melnikov A, Murugan A, Zhang X, *et al.*: **Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay.** *Nat Biotechnol.* 2012; **30**(3): 271–7.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kwasnieski JC, Fiore C, Chaudhari HG, *et al.*: **High-throughput functional testing of ENCODE segmentation predictions.** *Genome Res.* 2014; **24**(10): 1595–602.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Smith RP, Taher L, Patwardhan RP, *et al.*: **Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model.** *Nat Genet.* 2013; **45**(9): 1021–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Nguyen TA, Jones RD, Snaveley AR, *et al.*: **High-throughput functional comparison of promoter and enhancer activities.** *Genome Res.* 2016; **26**(8): 1023–33.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kheradpour P, Ernst J, Melnikov A, *et al.*: **Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay.** *Genome Res.* 2013; **23**(5): 800–11.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ernst J, Melnikov A, Zhang X, *et al.*: **Genome-scale high-resolution mapping of**





- activating and repressive nucleotides in regulatory regions. *Nat Biotechnol.* 2016; **34**(11): 1180–90.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. **F** Tewhey R, Kotliar D, Park DS, *et al.*: **Direct Identification of Hundreds of Expression-Modulating Variants using a Multiplexed Reporter Assay.** *Cell.* 2016; **165**(6): 1519–29.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
  15. **F** Ulirsch JC, Nandakumar SK, Wang L, *et al.*: **Systematic Functional Dissection of Common Genetic Variation Affecting Red Blood Cell Traits.** *Cell.* 2016; **165**(6): 1530–45.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
  16. **F** Arnold CD, Gerlach D, Stelzer C, *et al.*: **Genome-wide quantitative enhancer activity maps identified by STARR-seq.** *Science.* 2013; **339**(6123): 1074–7.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
  17. Muerdter F, Boryń LM, Arnold CD: **STARR-seq - principles and applications.** *Genomics.* 2015; **106**(3): 145–50.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  18. Yáñez-Cuna JO, Arnold CD, Stampfel G, *et al.*: **Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features.** *Genome Res.* 2014; **24**(7): 1147–56.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  19. Shlyueva D, Stelzer C, Gerlach D, *et al.*: **Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin.** *Mol Cell.* 2014; **54**(1): 180–92.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  20. Arnold CD, Gerlach D, Spies D, *et al.*: **Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution.** *Nat Genet.* 2014; **46**(7): 685–92.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  21. Zabidi MA, Arnold CD, Schernhuber K, *et al.*: **Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation.** *Nature.* 2015; **518**(7540): 556–9.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  22. Vanhille L, Griffon A, Maqbool MA, *et al.*: **High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq.** *Nat Commun.* 2015; **6**: 6905.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  23. Vockley CM, Guo C, Majoros WH, *et al.*: **Massively parallel quantification of the regulatory effects of noncoding genetic variation in a human cohort.** *Genome Res.* 2015; **25**(8): 1206–14.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  24. Shen SQ, Myers CA, Hughes AE, *et al.*: **Massively parallel cis-regulatory analysis in the mammalian central nervous system.** *Genome Res.* 2016; **26**(2): 238–55.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  25. Murtha M, Tokcaer-Keskin Z, Tang Z, *et al.*: **FIREWACH: high-throughput functional detection of transcriptional regulatory modules in mammalian cells.** *Nat Methods.* 2014; **11**(5): 559–65.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  26. Inoue F, Kircher M, Martin B, *et al.*: **A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity.** *Genome Res.* 2017; **27**(1): 38–52.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  27. **F** Ruf S, Symmons O, Uslu VV, *et al.*: **Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor.** *Nat Genet.* 2011; **43**(4): 379–86.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
  28. **F** Akhtar W, de Jong J, Pindyrin AV, *et al.*: **Chromatin position effects assayed by thousands of reporters integrated in parallel.** *Cell.* 2013; **154**(4): 914–27.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
  29. **F** Symmons O, Uslu VV, Tsujimura T, *et al.*: **Functional and topological characteristics of mammalian regulatory domains.** *Genome Res.* 2014; **24**(3): 390–400.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
  30. Dogan N, Wu W, Morrissey CS, *et al.*: **Occupancy by key transcription factors is a more accurate predictor of enhancer activity than histone modifications or chromatin accessibility.** *Epigenetics Chromatin.* 2015; **8**: 16.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  31. Bevington SL, Cauchy P, Cockerill PN: **Chromatin priming elements establish immunological memory in T cells without activating transcription: T cell memory is maintained by DNA elements which stably prime inducible genes without activating steady state transcription.** *Bioessays.* 2017; **39**(2).  
[PubMed Abstract](#) | [Publisher Full Text](#)
  32. **F** Engreitz JM, Haines JE, Perez EM, *et al.*: **Local regulation of gene expression by lncRNA promoters, transcription and splicing.** *Nature.* 2016; **539**(7629): 452–5.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
  33. Dao LTM, Galindo-Albarrán AO, Castro-Mondragon JA, *et al.*: **Genome-wide characterization of mammalian promoters with distal enhancer functions.** *Nat Genet.* 2017.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  34. Andersson R: **Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model.** *Bioessays.* 2015; **37**(3): 314–23.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  35. Pekowska A, Benoukraf T, Zacarias-Cabeza J, *et al.*: **H3K4 tri-methylation provides an epigenetic signature of active enhancers.** *EMBO J.* 2011; **30**(20): 4198–210.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  36. Koch F, Fenouil R, Gut M, *et al.*: **Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters.** *Nat Struct Mol Biol.* 2011; **18**(8): 956–63.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  37. Core LJ, Martins AL, Danko CG, *et al.*: **Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers.** *Nat Genet.* 2014; **46**(12): 1311–20.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  38. Diao Y, Fang R, Li B, *et al.*: **A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells.** *Nat Methods.* 2017; **14**(6): 629–35.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  39. Rajagopal N, Srinivasan S, Kooshesh K, *et al.*: **High-throughput mapping of regulatory DNA.** *Nat Biotechnol.* 2016; **34**(2): 167–74.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  40. Vierstra J, Reik A, Chang KH, *et al.*: **Functional footprinting of regulatory DNA.** *Nat Methods.* 2015; **12**(10): 927–30.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  41. Carver MC, Smith EC, Sher F, *et al.*: **BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis.** *Nature.* 2015; **527**(7577): 192–7.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  42. Sanjana NE, Wright J, Zheng K, *et al.*: **High-resolution interrogation of functional elements in the noncoding genome.** *Science.* 2016; **353**(6307): 1545–9.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  43. Diao Y, Li B, Meng Z, *et al.*: **A new class of temporarily phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening.** *Genome Res.* 2016; **26**(3): 397–405.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  44. Korkmaz G, Lopes R, Ugalde AP, *et al.*: **Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9.** *Nat Biotechnol.* 2016; **34**(2): 192–8.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  45. Hsu PD, Lander ES, Zhang F: **Development and applications of CRISPR-Cas9 for genome engineering.** *Cell.* 2014; **157**(6): 1262–78.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  46. Klann TS, Black JB, Chellappan M, *et al.*: **CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome.** *Nat Biotechnol.* 2017; **35**(6): 561–568.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  47. **F** Fulco CP, Munschauer M, Anyoha R, *et al.*: **Systematic mapping of functional enhancer-promoter connections with CRISPR interference.** *Science.* 2016; **354**(6313): 769–73.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
  48. Xie S, Duan J, Li B, *et al.*: **Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells.** *Mol Cell.* 2017; **66**(2): 285–299.e5.  
[PubMed Abstract](#) | [Publisher Full Text](#)
  49. Sati S, Cavalli G: **Chromosome conformation capture technologies and their impact in understanding genome function.** *Chromosoma.* 2017; **126**(1): 33–44.  
[PubMed Abstract](#) | [Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status:   

---

## Editorial Note on the Review Process

**Faculty Reviews** are review articles written by the prestigious Members of **Faculty Opinions**. The articles are commissioned and peer reviewed before publication to ensure that the final, published version is comprehensive and accessible. The reviewers who approved the final version are listed with their names and affiliations.

---

## The reviewers who approved this article are:

### Version 1

- Bart Deplancke**  
Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland  
**Competing Interests:** No competing interests were disclosed.
- Douglas Vernimmen**  
The Roslin Institute, University of Edinburgh, Edinburgh, UK  
**Competing Interests:** No competing interests were disclosed.
- Alvaro Rada-Iglesias**  
Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany  
**Competing Interests:** No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**