Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Applications of high-throughput reporter assays to gene regulation studies

Benedetta D'Elia[1,4] and Juan Fuxman Bass[1,2,3,4]

Determining the rules of transcriptional regulation, associated with a complex transcription factor grammar, is fundamental to understand the control of most biological processes, disease mechanisms, and evolution. High-throughput reporter assays, such as MPRAs and STARR-seq, have enabled systematic functional annotations of genomes and have provided large substrates for training machine learning models to determine these rules, predict the activity of native and synthetic elements, and design elements for different applications. This review provides an overview of high-throughput reporter assays and their applications for the study of enhancers, promoters, silencers and insulators. We discuss how these assays help identify causal disease-associated non-coding variants and design synthetic elements with desired features for functional studies or therapeutic purposes.

Addresses
[1] Department of Molecular Biology, Cell Biology & Biochemistry, Boston University, Boston, MA, USA
[2] Department of Biology, Boston University, Boston, MA, USA
[3] Bioinformatics Program, Boston University, Boston, MA, USA
[4] Biological Design Center, Boston University, Boston, MA, USA

Corresponding author: Fuxman Bass, Juan (fuxman@bu.edu)

Current Opinion in Structural Biology 2025, **94**:103105

This review comes from a themed issue on **Protein Nucleic Acid Interactions (2025)**

Edited by **Shandar Ahmad** and **Elodie Laine**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.sbi.2025.103105

Introduction

Gene expression programs are controlled by the binding of transcription factors (TFs) to *cis*-regulatory elements (CREs) such as promoters, enhancers, and silencers [1]. Understanding the logic and grammar of these binding events is fundamental to determine mechanisms of gene regulation and their role in development, disease, and evolution. This requires multi-pronged studies to identify CREs active in particular conditions, determine the TF motifs contributing to activity, identify the combinatorial mechanisms involved in regulation, and the rational design of CREs with specific functionalities for bioengineering applications or to test different hypotheses of regulatory mechanisms [2]. Although these have been central, long-standing goals in the gene regulation field, they remained mostly elusive for several decades given the large number of sequences that often need to be evaluated.

With the advent of high-throughput sequencing, several complementary approaches have been developed to fill this gap. For example, chromatin immunoprecipitation followed by sequencing (ChIP-seq) [3], DNase-seq [4], and the assay for transposase-accessible chromatin with sequencing (ATAC-seq) [5], have identified regions bound by specific TFs and cofactors, open chromatin, and different chromatin states. These studies have also provided insights into the DNA binding logic of hundreds of thousands of CREs based on combinations of TF binding peaks or footprints [6–8]. However, these studies do not directly measure transcriptional activity and rely on the natural DNA sequence variability present in the studied genomes. Given that this natural variability only represents a very small fraction of the possible sequence space, these studies are often limited in the hypotheses that can be tested regarding regulatory logic, effect of variants, and design of novel sequences [9].

In the last 15 years, high-throughput reporter assays have emerged as major tools to directly measure putative CRE activity across thousands of sequences simultaneously. These are an adaptation of reporter assays that became increasingly popular in the 80s and 90s, which couple candidate regulatory sequences to reporter genes and unique sequence barcodes, enabling the measurement of transcriptional activity in a quantitative, highly reproducible manner, across tens of thousands of sequences. A pioneering 2009 study laid the ground through a saturation mutagenesis-based in vitro assay to analyze the effect of mutations at every position in bacteriophage and mammalian core promoters, thereby providing a powerful tool for the annotation of functional regulatory elements [10]. Since then, different high-throughput reporter assays, such as massively parallel reporter assays (MPRAs) and Self-Transcribing Active Regulatory Region sequencing (STARR-seq), [10–14], have been developed to identify and dissect the mechanism of regulation of

thousands of CREs, determine the effects of genetic variants, train statistical models that provide a better understanding of regulatory mechanisms, and to rationally design synthetic elements sequences. Given the breadth of this field, it is impossible to mention all the important work performed. This review will therefore provide a broad view of the main experimental approaches currently used as well as their different applications to CRE studies.

High-throughput reporter assay methodologies

Different high-throughput reporter assays evaluate the activity of thousands to millions of candidate CREs simultaneously using a variety of construct designs depending on the types of CREs evaluated and the goals of the studies. The main distinctive features include: 1) the source of DNA for library construction (sheared genomic DNA or synthesized oligonucleotide pools); 2) the cloning position of the DNA sequences tested (upstream or downstream of the transcription start); 3) the library delivery method into cells or living organisms; and 4) the type of readout and analyses (barcodes or peaks). For instance, some assays employ synthesized DNA fragments representing candidate CREs, allowing for full control to test specific hypothesis about the impact of variants, perturbation of TF motifs, or specific arrangement of TF binding sites, while others rely on sheared genomic DNA, capturing a broader spectrum of native sequences and exploiting natural variation in the genome. Understanding these design variations and their applications is crucial for selecting the appropriate methodology for different research projects. Here, we describe and compare two major classes of high-throughput reporter assay methodologies, MPRAs and STARR-seq, and their variations (Table 1).

MPRAs

MPRAs test thousands of DNA sequences synthesized as oligonucleotide pools and transcriptional activity is measured using barcodes placed in the transcribed region (Figure 1a). Episomal MPRAs rely on transient expression systems, where plasmid libraries are transfected into cells, commonly via electroporation. Some key advantages of episomal MPRAs are their rapid turnaround time, as they bypass the need for genomic integration, and the high dynamic range in activity detected. This makes them particularly useful for high-throughput screening of regulatory sequences in diverse cellular contexts. However, episomal MPRAs present some limitations: sequences tested are usually shorter than most native CREs, cells should be easily transfectable, and they may not fully recapitulate the chromatinized genome environment, potentially affecting the physiological relevance of the results [15]. Despite these limitations, episomal MPRAs remain widely used for dissecting enhancer and promoter activity due to their scalability and efficiency. Pioneering studies have laid the ground for using MPRAs to discover gene regulatory mechanisms and demonstrated their utility in mapping functional motifs that contribute to activity across cell types using saturation mutagenesis, where single nucleotide mutations are introduced at each position of the target sequence [10,12,14].

lentiMPRAs have been developed to bypass limitations of episomal MPRAs by employing lentiviral transduction to stably integrate candidate CREs into the genome. This approach allows for long-term expression, useful to evaluate influence of chromatin state, epigenetic modifications, or slower transcriptional processes [16,17]. Additionally, lentiviruses can infect a broad range of cell types. However, a key challenge is the random integration of lentiviruses which can introduce positional effects that influence reporter expression. While efforts have been made to mitigate these effects through bulk analysis of multiple integration events, variability, efficiency and selection biases remain a concern [15]. Despite these limitations, several studies have highlighted lentiMPRA utility in identifying functional enhancers and understanding gene regulation mechanisms in a chromatin context [16–19].

Bioinformatic pipelines to analyse MPRA data usually include barcode-element mapping, barcode counting, RNA to DNA count normalization, filtering and differential activity analysis through tools such as DESeq2, and quality controls. While not as standardized as other genomic assays analyses, multiple analysis pipelines have been developed, including MPRAflow, mpralm, MPRAanalyze and Sharpr-MPRA [17,20–22]. Downstream analyses, however, often vary according to the experimental design, including the type of elements tested (genomic or synthetic design), delivery method (episomal vs lentiviral), and cell type or condition, and use *ad hoc* pipelines that consider their specific goals. Due to the use of multiple barcodes per sequence, MPRAs are highly reproducible (0.97–0.99 correlation between biological replicates) and highly sensitive, enabling the detection of low active CREs and differential activity between CRE alleles as low as 10 % [14,23].

STARR-seq

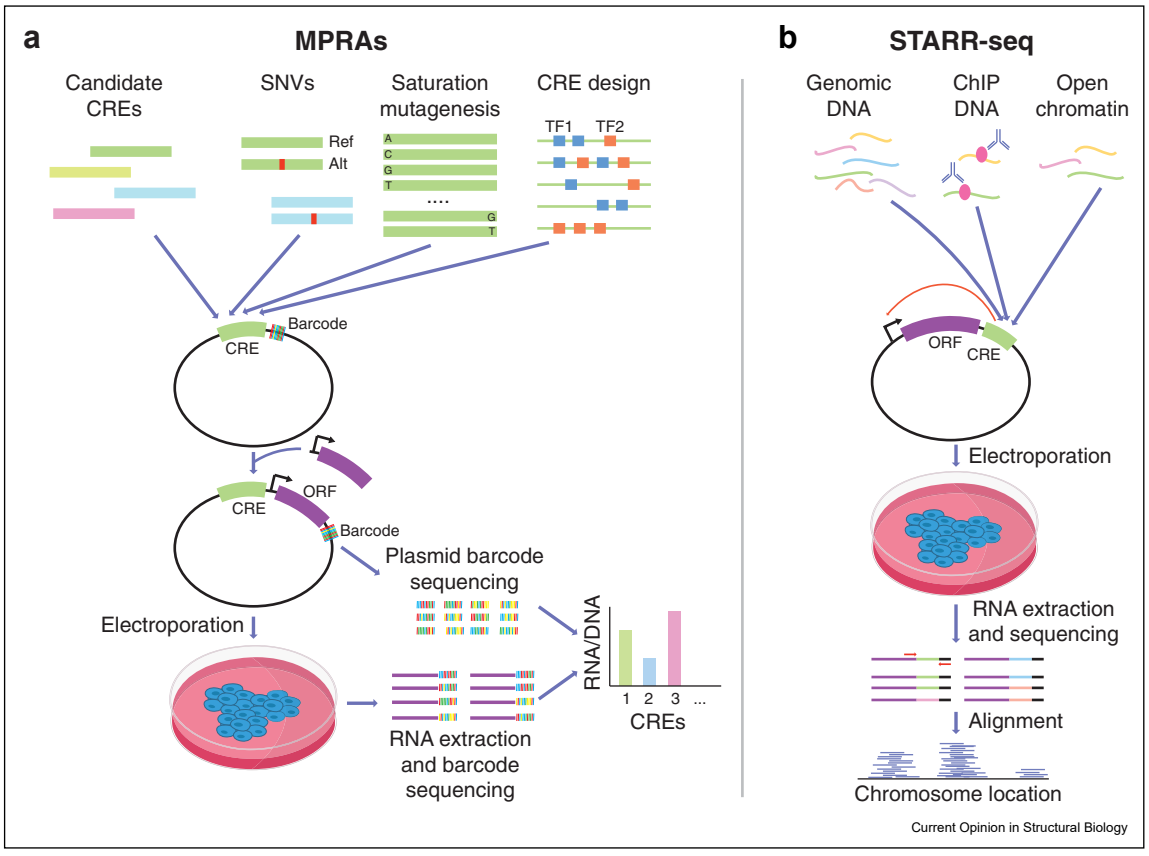
Unlike MPRAs, which rely on determining barcode read counts, STARR-seq directly places candidate CREs within the 3′ UTR of the construct (Figure 1b). This enables measuring the activity of enhancers (or silencers), which function independently of their position, thus driving their own transcription. PCR around the cloning site in the cDNA fraction, followed by sequencing and mapping to genomic sites, allows to identify peaks of high activity relative to the original plasmid pool, genome-wide [11]. This is because

Table 1

An overview of different high-throughput reporter assay methodologies.

Method	Type of DNA	Barcodes	Position of DNA in construct	Type of promoter	Goal	Sources
Episomal MPRA	Synthesized	3' UTR	Upstream of promoter	Minimal	Test CRE activity, impact of SNVs, and synthetic CREs	[12–14]
Saturation-mutagenesis MPRA	Synthesized	3' UTR	Upstream of promoter	Minimal	Test impact of CRE mutations	[12,13,51]
lentiMPRA	Synthesized	5' UTR	Upstream of promoter	Minimal	Test activity of integrated CREs	[16–18]
STARR-seq	Sheared, genomic	No	Downstream of promoter	Minimal	Identify enhancers genome-wide	[11,35]
STAP-seq	Sheared, genomic	5' UTR	Upstream of promoter	Minimal	Determine responsiveness of genomic sequences to enhancers	[37]
SuRE-seq	Sheared, genomic	5' UTR	Upstream of ORF	Promoterless	Identify enhancers and promoters genome-wide	[31,32]
ChIP-STARR-seq	ChIP fragments	No	Downstream of promoter	Minimal	Test TF-bound sequences for activity	[27]
ATAC-STARR-seq	Genomic/tagmented	No	Downstream of promoter	Minimal	Test open chromatin regions for activity	[24,28]
TSS-MPRA	Synthetic	5' UTR	Upstream of reporter	Promoterless	Measures TSS profiles	[36]
Silencer-seq	Sheared, genomic	No	Downstream of promoter	Strong enhancer	Identify silencers genome-wide	[40]
ssSTARR-seq	Sheared, genomic	No	Downstream of promoter	Constitutive	Identify silencers genome-wide	[42]
MPRAduo	Synthesized	dual: coupled to CRE and silencer libraries	Upstream of promoter	Minimal	Test enhancer/silencer combinations	[41]
MPIRE	Synthesized	dual: genomic BC and *cis*-reg BC	Upstream of promoter	Minimal	Identify insulators genome-wide	[44]
snpSTARR-seq	Sheared, pooled, genomic	UMIs	3' UTR	Minimal	Determine SNV effects	[54]
scMPRA	Synthesized	dual: candidate and random BC	Upstream of promoter	Minimal	Test CREs activity in single cells	[70,71]

Figure 1



Overview of MPRA and STARR-seq workflows. a. In MPRAs, candidate DNA sequences synthesized as oligonucleotide pools are associated with random, unique barcodes either during synthesis or by PCR, and cloned into a plasmid. Sequences include putative CREs, sequences containing genetic variants or systematic sequence perturbations, and engineered sequences containing TF motifs. Sequencing of this construct library enables matching barcodes to the synthesized sequences to establish a barcode directory. A minimal promoter and an open reading frame (ORF) are cloned into this initial library between the library DNA and the barcodes so that the library is located upstream of the transcription start site and the barcode is placed in the 3'UTR. The construct is then introduced into cells of interest via transfection, electroporation, or through lentiviral infection, depending on the application. RNA is then extracted from cells, reverse transcribed to cDNA, and the region surrounding the barcodes are amplified by PCR and sequenced to determine barcode read counts. Finally, barcode read counts are determined in the cDNA fraction relative to those in the plasmid fraction to determine transcriptional activity of the candidate CRE. b. In STARR-seq, sheared genomic, ChIP- or ATAC-DNA can be cloned into a vector into the 3' UTR, containing a reporter gene, and downstream of a promoter, where an active fragment can drive transcription of the reporter gene. After delivery into cells through transfection, RNA is extracted and the transcripts sequenced to determine which regulatory elements are associated with high levels of transcription via peak calling.

millions of genomic DNA fragments can be cloned and tested, without the need to synthesize massive libraries, which could be cost-prohibitive. Indeed, STARR-seq has been employed to dissect the functional architecture of enhancers across diverse genomic landscapes [11,24–26]. Additionally, STARR-seq can be performed on DNA sequences that have been immunoprecipitated from chromatin or derived from open chromatin regions (e.g., ATAC libraries), thereby focusing studies on functional elements [27,28]. Because STARR-seq is conducted using episomal vectors, context dependency remains a limitation, similar to episomal MPRAs [29,30]. Variations of STARR-seq, such as SuRE-seq, detect regions with autonomous transcriptional activity and improve CRE characterization [31,32]. In SuRE, random genomic DNA fragments are cloned in a promoterless barcoded reporter plasmid, combining strengths of both STARR-seq and MPRAs, to identify fragments containing transcription start site (TSS).

Bioinformatic analysis of STARR-seq data usually involves customizable pipelines that include raw data preprocessing, mapping input DNA and output RNA reads to the reference genome, peak calling, and peak annotation. Although initial studies used peak calling algorithms designed for ChIP-seq (e.g., MACS2), STARR-seq-specific peak calling algorithms, such as STARRPeaker [33] and CRADLE [34], were developed that account for sequence-associated biases. While STARR-seq peaks have shown consistency between

biological replicates [26], the absence of barcodes in the traditional design, as well as the direct quantification of enhancer-driven transcripts can all influence overall reproducibility. For this reason, Das et al. compiled a useful resource of STARR-seq experiments and provides recommendations to increase reproducibility, including intermediate sample validations and quality control assessments [35].

Identification and characterization of *cis*-regulatory elements

High-throughput reporter assays have been instrumental to identify different types of genomic elements, including gene promoters, enhancers, silencers, and insulators, as well as to determine the motifs that contribute to their function (Figure 2a). These approaches either exploit the natural variation occurring within genomes or employ mutagenized sequences to investigate the impact of different TF motifs on activity.

MPRAs have revealed different promoter architectures and mapped the sequence features that dictate strength, specificity, and activity [36]. SuRE-seq experiments mapped autonomous promoter activity, genome-wide [32]. These and other studies showed that TSS-proximal regions are responsible for most autonomous promoter activity, that many promoters overlap with CpG islands, that enhancers often act as autonomous promoters, and identified sequence features using generalized linear models. Further, lentiMPRAs have expanded these results by showing that promoters display a preference for strand directionality and that their 200-nucleotide cores function as 'on switches' that are generally not cell type specific, contrary to enhancers [16].

Genome-wide STARR-seq has generated comprehensive enhancer maps, revealing thousands of enhancers with both constitutive and condition-specific activities across cell types and organisms [11,26,29,37]. Episomal and lentiMPRA experiments have shown that enhancers have weaker orientation bias and increased tissue and cell type specificity compared to promoter elements [16]. Saturation or tiling mutagenesis approaches identified sequence motifs that positively and negatively contribute to enhancer activity and cell type specificity [12,13,22,38]. Testing libraries of tens to hundreds of thousands of enhancer variants revealed that most enhancers are robust to most single nucleotide substitutions, while substitutions in multiple TF binding sites can additively affect enhancer activity [12,13].

Silencers have also been identified and characterized using high-throughput reporter assays. One study used DNase hypersensitivity data to exclude chromatin and binding signatures associated with activating CREs, to then build a STARR-seq library for testing "uncharacterized elements" hypothesized to be silencers [39]. 40 % of these putative elements functioned as silencers, and motif enrichment analysis confirmed that silencers harbor motifs for the binding of canonical transcriptional repressors, such as REST, JUN and NRF2. Another group developed Silencer-seq, a STARR-seq variation using a strong promoter/enhancer to detect repressive activity, which identified ~1000 silencers in *Drosophila* S2 cells [40]. This study uncovered novel silencer features: they are rarely found in open chromatin or heterochromatin regions; their TFs binding sites have distinct nucleosome footprints that render them undetectable by other techniques such as ATAC-seq or DNase-seq; and that silencers can act as enhancers or insulators in different contexts or cell-types [24,39–42]. This is consistent with an *in vivo* study in *Drosophila* that showed that silencers can function as enhancers in different tissues [43].

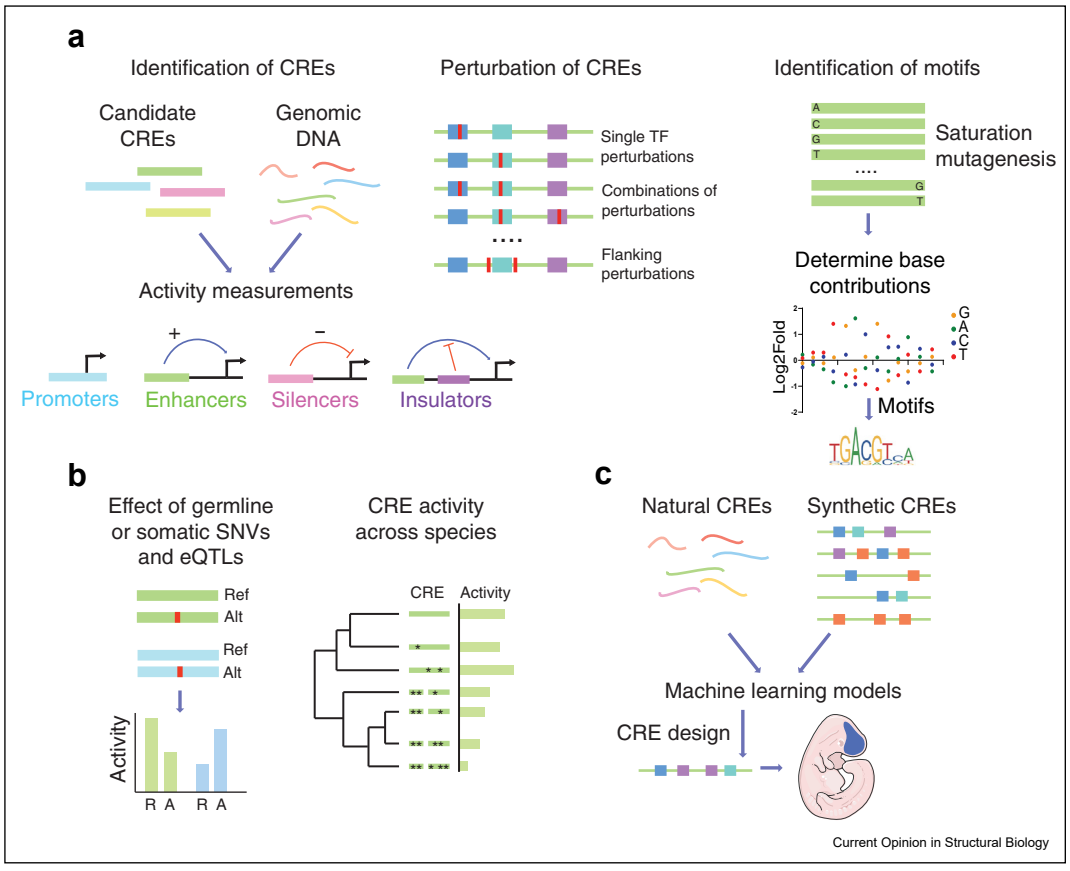
Finally, insulators, which function as boundary elements that prevent interactions between enhancers and promoters, have proven more challenging to study, as their activity is highly context-dependent. A recent study using multiple landing pads randomly integrated into the genome determined insulator capacity genome-wide and found that insulator activity is highly context dependent and dependent on CTCF or B-box motifs [44].

Determination of transcriptional effects of genetic variants

Sequence variation in CREs can have profound effects on gene expression often leading to disease [1,45]. Several Mendelian diseases have been associated with promoter or enhancer mutations, which have been characterized using low-throughput reporter, DNA binding, and *in vivo* assays. Genome-wide association studies (GWAS) have also identified hundreds of loci associated with complex diseases such as psychiatric disorders, cardiovascular diseases, autoimmune diseases, and obesity. Multiple studies have shown that most GWAS variants likely affect gene expression rather than protein coding sequences [46]. However, because of linkage disequilibrium, GWAS cannot pinpoint causal variants within the statistically associated loci, which are often in the order of tens to hundreds [47]. Therefore, this requires evaluating thousands of variants across loci to determine their impact on transcriptional activity. Expression quantitative trait loci (eQTLs) have similar limitations as associations cannot distinguish closely linked variants.

High-throughput reporter assays have been instrumental to characterize genetic variants in disease, development, and evolution (Figure 2b). By synthesizing libraries with thousands of variants, MPRAs can functionally assess how CRE changes affect transcriptional output. For instance, MPRAs have identified

Figure 2



Applications of high-throughput reporter assays. **(a)**. Use of MPRAs and STARR-seq to identify different types of CREs (promoters, enhancers, silencers and insulators). Identification of TF motifs that contribute to CRE activity can be performed by specific single- or multi- TF binding site perturbations or by saturation mutagenesis to determine base contributions at a single-nucleotide resolution. **(b)**. High-throughput reporter assays can help identify germline and somatic SNVs and eQTLs that affect transcriptional activity and are associated disorders such as cancer and genetic diseases, and study the evolutionary variation in CRE activity across species. **(c)**. Deciphering regulatory grammar through MPRAs and STARR-seq datasets enables the prediction of de novo sequences via machine learning models, which instruct the design of tissue-specific synthetic elements for therapeutic applications and fine-tuned gene expression.

hundreds of causal alleles from large candidate eQTL and GWAS loci [14,48]. Similarly, SuRE-seq of random genomic fragments from four individuals assessed 6 million single nucleotide variants (SNVs) leading to the identification of thousands of SNVs affecting transcriptional activity and helped identify functional hits within GWAS and eQTL loci [31]. These are just a few of an increasing set of studies identifying germline and somatic variants associated with a wide range of diseases including autoimmune diseases, neurological disorders, and cancer [49–56].

High-throughput reporter assays have enormous potential for applications in fields such as evolutionary biology [45]. MPRAs have been employed to characterize liver enhancer evolution in primates, where 84 SNVs were found to correlate with functional changes such as CpG deamination [57]. Another study explored the functional consequences of the modern humans/archaic hominins split and showed that 23% of CREs drove differential activity between species [58]. These studies are beginning to expand evolutionary studies across different model and non-model organisms.

Elucidating and exploiting the rules of transcriptional regulation

Understanding the regulatory grammar of gene expression is central to determining when, where, and how genes are expressed and to predict the effect of disease variants from sequence. In addition, this will enable the rational design of synthetic CREs with specific features to answer mechanistic questions and for the development of therapeutic tools for cell type- or condition-specific expression. Given the number of potential TF binding site combinations, orientations and spacings in CREs, studies to understand these grammar rules require tens of thousands of sequences to train machine learning models that can make accurate predictions.

High-throughput reporter assays provide a large substrate for model training by testing sequences that exploit natural sequence variation within genomes or synthetic elements with user-selected TF binding site architectures (Figure 2c). MPRAs have also been used to test some of these model predictions, including the design of de novo synthetic elements. For instance, a study tested a lentiMPRA library of ~200,000 sequences containing homotypic and heterotypic combinations of 18 liver TF binding sites [59]. This and other studies found multiple features that impact regulatory activity, including number, orientation, distance, and order of TF binding sites from the TSS, as well as chromatin context [2,60,61].

Machine learning models trained on MPRA data have provided accurate activity predictions of mammalian, bacterial, and *Drosophila* enhancers and promoters [61–67]. For instance, a cell-type specific deep learning model trained on MPRAs was used to predict human promoter activity genome-wide and design synthetic human promoters that did not resemble any human genome sequence, suggesting that the model learned regulatory rules de novo [64]. Another study developed DeepSTARR, a deep learning model to predict enhancer activity from *Drosophila* sequences using STARR-seq data [62]. This model learned sequence regulatory features such as TF motifs, orientation, effect of flanking sequences, and motif cooperativity, which enabled the design of strong synthetic developmental enhancers with specific motif composition and syntax. A follow up study optimized the DeepSTARR architecture and employed scATAC-seq and *in vivo* enhancer activity datasets to design tissue-specific enhancers for five *Drosophila* embryo tissues [65]. Another study leveraged a deep neural network model termed Malinois trained on MPRA data to generate synthetic sequences with increased cell-type specificity relative to native sequences and that even employs unique TF combinations not found in the human genome [66]. As the scale and diversity of MPRA experiments continue to grow, centralized databases like MPRAbase offer a valuable resource for the storage of MPRA results across different organisms, cell types, and experimental conditions, facilitating the training of machine learning models aimed at deciphering, as well as engineering, cis-regulatory logic [67]. These advances in synthetic CRE design have broad implications for biotechnology, synthetic biology, and therapeutic applications, which require fine-tuned gene expression control, and may reduce the need for in vitro and *in vivo* validation.

Limitations of high-throughput reporter assays

Throughout this review, we highlighted the advantages of high-throughput reporter assays to identify and characterize CREs; however, there are some considerations and limitations for their use in gene regulation experiments. A major challenge for both MPRAs and STARR-seq is their reliance on testing sequences outside of their native chromatin context. Because the reporter constructs tested are episomal, randomly integrated, or integrated in landing pads, they fail to recapitulate native chromatin features, such as histone modifications, nucleosome positioning, or 3D interactions between CREs. This limits their ability to capture context-dependent regulation, particularly for elements that rely on specific locus architecture for activity. Conversely, other genomic assays such as ChIP-seq and ATAC-seq can profile TF or cofactor binding, histone modifications, or accessibility of native chromatin, but do not directly measure regulatory function. Another challenge of high-throughput reporter assays, shared by ChIP-seq and ATAC-seq approaches, involves assigning CREs to specific target genes, especially for elements such as distal enhancers. More recent approaches, such as CRISPR interference tiling across genomic loci are beginning to enable endogenous perturbations of CREs and can address the contextual limitations and CRE-target assignments [68,69]. However, these assays are experimentally intensive and lack the versatility of MPRAs to assay synthetic sequences or multiple loci in parallel. Ultimately, a combination of different complementary approaches is need dissect CREs in context-specific manner.

Future outlooks

Recent advances in single-cell methods, *in vivo* systems, and assays exploring gene–environment interactions, as well as machine learning models, are expanding the horizons of functional genomics. High-throughput reporter assays are generally conducted using populations of millions of cells, which may mask the contributions of individual cells and often lack the complexity of tissue environments. The advent of single-cell MPRAs (scMPRAs) addresses this limitation, revealing the diversity of CRE activity across different cell types and states [70,71].

Further, extending MPRAs to *in vivo* systems would allow assessment of CRE activity and specificity within their native cellular types, states, and physiological conditions. Pioneering studies performed MPRAs in explanted newborn mouse retinas to identify retina active CREs [72,73]. More recent studies have performed whole-animal *in vivo* MPRAs, where intravenous injection of plasmid or adeno-associated virus MPRA libraries to mouse tissues lead to the identification of brain- and liver-specific enhancers and the effects of synthetic and native TF binding site disruption on enhancer activity [74,75]. Expanding these studies to other sets of elements and tissues would offer a powerful approach to dissect the contributions of individual cells to CRE activity and effect of genetic variants in different tissues or disease contexts.

Another major area of future development concerns "Response MPRAs" [76], aimed at studying how CREs respond to environmental stimuli. A few groups have already used this growing approach to study CRE activation/repression by ligands. For example, STARR-seq identified CREs that respond to the glucocorticoid dexamethasone, as well as allele-specific effects [27,77]. Another study showed that retinoic acid, implicated with risk for major depressive disorder, increased regulatory activity of allelic variants harbouring retinoic receptor motifs [78]. Future work is expected to extend these studies to shed light onto the mechanisms governing CRE activity in response to drugs, hormones, metabolites, as well as endogenous and exogenous immune stimuli and stresses. Coupling Response MPRAs with machine learning models will also expand our ability to predict and design stimulus-specific CREs. Along these lines, a recent study designed a library of synthetic promoters that serve as readouts of cellular stimuli [79]. Another study used a similar approach to design reporters that respond to the activation by specific TFs [80]. The integration of machine learning models and experimental MPRA datasets in different conditions, cell types, and *in vivo* settings will lead to more precise models of gene regulation, advancing our ability to predict causal variants, design synthetic reporters, and develop cell type- and context-specific elements for biotechnology or disease treatment applications.

Author's contributions

B.D. and J.F.B. wrote and edited the manuscript.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Juan Fuxman Bass reports financial support was provided by National Institutes of Health. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by National Institutes of Health grants R35GM128625 and R01HD104971.

Data availability

No data was used for the research described in the article.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Chatterjee S, Ahituv N: **Gene regulatory elements, major drivers of human disease**. *Annu Rev Genom Hum Genet* 2017, **18**:45–63, https://doi.org/10.1146/annurev-genom-091416-035537.

2. Grossman SR, Zhang X, Wang L, Engreitz J, Melnikov A, Rogov P, *et al.*: **Systematic dissection of genomic features determining transcription factor binding and enhancer function**. *Proc Natl Acad Sci USA* 2017, **114**, https://doi.org/10.1073/pnas.1621150114.

3. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, *et al.*: **Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing**. *Nat Methods* 2007, **4**:651–657, https://doi.org/10.1038/nmeth1068.

4. Song L, Crawford GE: **DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells**. *Cold Spring Harb Protoc* 2010, **2010**, https://doi.org/10.1101/pdb.prot5384. pdb.prot5384.

5. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ: **Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position**. *Nat Methods* 2013, **10**:1213–1218, https://doi.org/10.1038/nmeth.2688.

6. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, *et al.*: **The NIH roadmap epigenomics mapping consortium**. *Nat Biotechnol* 2010, **28**:1045–1048, https://doi.org/10.1038/nbt1010-1045.

7. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, *et al.*: **An integrated encyclopedia of DNA elements in the human genome**. *Nature* 2012, **489**:57–74, https://doi.org/10.1038/nature11247.

8. ENCODE Project Consortium: **A user's guide to the encyclopedia of DNA elements (ENCODE)**. *PLoS Biol* 2011, **9**, e1001046, https://doi.org/10.1371/journal.pbio.1001046.

9. White MA: **Understanding how cis-regulatory function is encoded in DNA sequence using massively parallel reporter assays and designed sequences**. *Genomics (San Diego, Calif)* 2015, **106**:165–170, https://doi.org/10.1016/j.ygeno.2015.06.003.

10. Patwardhan RP, Lee C, Litvin O, Young DL, Pe'er D, Shendure J: **High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis**. *Nat Biotechnol* 2009, **27**:1173–1175, https://doi.org/10.1038/nbt.1589.

11. Arnold CD, Gerlach D, Stelzer C, Boryń ŁM, Rath M, Stark A: **Genome-wide quantitative enhancer activity maps identified by STARR-seq**. *Science* 2013, **339**:1074–1077, https://doi.org/10.1126/science.1232542.

12. Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, *et al.*: **Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay**. *Nat Biotechnol* 2012, **30**:271–277, https://doi.org/10.1038/nbt.2137.

13. Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, *et al.*: **Massively parallel functional dissection of mammalian enhancers in vivo**. *Nat Biotechnol* 2012, **30**:265–270, https://doi.org/10.1038/nbt.2136.

14. Tewhey R, Kotliar D, Park DS, Liu B, Winnicki S, Reilly SK, *et al.*: **Direct identification of hundreds of expression-modulating variants using a multiplexed reporter assay**. *Cell* 2016, **165**:1519–1529, https://doi.org/10.1016/j.cell.2016.04.027.

15. Klein JC, Agarwal V, Inoue F, Keith A, Martin B, Kircher M, *et al.*: **A systematic evaluation of the design and context dependencies of massively parallel reporter assays**. *Nat Methods* 2020, **17**:1083–1091, https://doi.org/10.1038/s41592-020-0965-y.

16. ** Agarwal V, Inoue F, Schubach M, Penzar D, Martin BK, Dash PM, *et al.*: **Massively parallel characterization of transcriptional regulatory elements**. *Nature* 2025:1–10, https://doi.org/10.1038/s41586-024-08430-9.
This paper used lentiMPRAs to test the regulatory activity of more than 680,000 sequences and found sequence features that control the activity of promoters and enhancers.

17. Gordon MG, Inoue F, Martin B, Schubach M, Agarwal V, Whalen S, *et al.*: **lentiMPRA and MPRAflow for high-**

throughput functional characterization of gene regulatory elements. *Nat Protoc* 2020, **15**:2387–2412, https://doi.org/10.1038/s41596-020-0333-5.

18. Inoue F, Kircher M, Martin B, Cooper GM, Witten DM, McManus MT, *et al.*: **A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity**. *Genome Res* 2017, **27**:38–52, https://doi.org/10.1101/gr.212092.116.

19. Maricque BB, Dougherty JD, Cohen BA: **A genome-integrated massively parallel reporter assay reveals DNA sequence determinants of cis-regulatory activity in neural cells**. *Nucleic Acids Res* 2017, **45**:e16, https://doi.org/10.1093/nar/gkw942.

20. Myint L, Avramopoulos DG, Goff LA, Hansen KD: **Linear models enable powerful differential activity analysis in massively parallel reporter assays**. *BMC Genom* 2019, **20**:209, https://doi.org/10.1186/s12864-019-5556-x.

21. Ashuach T, Fischer DS, Kreimer A, Ahituv N, Theis FJ, Yosef N: **MPRAnalyze: statistical framework for massively parallel reporter assays**. *Genome Biol* 2019, **20**:183, https://doi.org/10.1186/s13059-019-1787-z.

22. Ernst J, Melnikov A, Zhang X, Wang L, Rogov P, Mikkelsen TS, *et al.*: **Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions**. *Nat Biotechnol* 2016, **34**:1180–1190, https://doi.org/10.1038/nbt.3678.

23. Kreimer A, Ashuach T, Inoue F, Khodaverdian A, Deng C, Yosef N, *et al.*: **Massively parallel reporter perturbation assays uncover temporal regulatory architecture during neural differentiation**. *Nat Commun* 2022, **13**:1504, https://doi.org/10.1038/s41467-022-28659-0.

24. Hansen TJ, Hodges E: **ATAC-STARR-seq reveals transcription factor–bound activators and silencers within chromatin-accessible regions of the human genome**. *Genome Res* 2022, **32**:1529–1541, https://doi.org/10.1101/gr.276766.122.

25. Neumayr C, Pagani M, Stark A, Arnold CD: **STARR-Seq and UMI-STARR-seq: assessing enhancer activities for genome-wide-, high-, and low-complexity candidate libraries**. *Curr Protoc Mol Biol* 2019, **128**, e105, https://doi.org/10.1002/cpmb.105.

26. Peng T, Zhai Y, Atlasi Y, ter Huurne M, Marks H, Stunnenberg HG, *et al.*: **STARR-seq identifies active, chromatin-masked, and dormant enhancers in pluripotent mouse embryonic stem cells**. *Genome Biol* 2020, **21**:243, https://doi.org/10.1186/s13059-020-02156-3.

27. Vockley CM, D'Ippolito AM, McDowell IC, Majoros WH, Safi A, Song L, *et al.*: **Direct GR binding sites potentiate clusters of TF binding across the human genome**. *Cell* 2016, **166**:1269–1281.e19, https://doi.org/10.1016/j.cell.2016.07.049.

28. Wang X, He L, Goggin SM, Saadat A, Wang L, Sinnott-Armstrong N, *et al.*: **High-resolution genome-wide functional dissection of transcriptional regulatory regions and nucleotides in human**. *Nat Commun* 2018, **9**:5380, https://doi.org/10.1038/s41467-018-07746-1.

29. Muerdter F, Boryń ŁM, Woodfin AR, Neumayr C, Rath M, Zabidi MA, *et al.*: **Resolving systematic errors in widely used enhancer activity assays in human cells**. *Nat Methods* 2018, **15**:141–149, https://doi.org/10.1038/nmeth.4534.

30. Shlyueva D, Stampfel G, Stark A: **Transcriptional enhancers: from properties to genome-wide predictions**. *Nat Rev Genet* 2014, **15**:272–286, https://doi.org/10.1038/nrg3682.

31. van Arensbergen J, Pagie L, FitzPatrick VD, de Haas M, Baltissen MP, Comoglio F, *et al.*: **High-throughput identification of human SNPs affecting regulatory element activity**. *Nat Genet* 2019, **51**:1160–1169, https://doi.org/10.1038/s41588-019-0455-2.

32. van Arensbergen J, FitzPatrick VD, de Haas M, Pagie L, Sluimer J, Bussemaker HJ, *et al.*: **Genome-wide mapping of autonomous promoter activity in human cells**. *Nat Biotechnol* 2017, **35**:145–153, https://doi.org/10.1038/nbt.3754.

33. Lee D, Shi M, Moran J, Wall M, Zhang J, Liu J, *et al.*: **STARRPeaker: uniform processing and accurate identification of STARR-seq active regions**. *Genome Biol* 2020, **21**:298, https://doi.org/10.1186/s13059-020-02194-x.

34. Kim Y-S, Johnson GD, Seo J, Barrera A, Cowart TN, Majoros WH, *et al.*: **Correcting signal biases and detecting regulatory elements in STARR-seq data**. *Genome Res* 2021, **31**:877–889, https://doi.org/10.1101/gr.269209.120.

35. Das M, Hossain A, Banerjee D, Praul CA, Girirajan S: **Challenges and considerations for reproducibility of STARR-seq assays**. *Genome Res* 2023, **33**:479–495, https://doi.org/10.1101/gr.277204.122.

36. Guzman C, Duttke S, Zhu Y, De Arruda Saldanha C, Downes NL, Benner C, *et al.*: **Combining TSS-MPRA and sensitive TSS profile dissimilarity scoring to study the sequence determinants of transcription initiation**. *Nucleic Acids Res* 2023, **51**, e80, https://doi.org/10.1093/nar/gkad562.

37. Arnold CD, Zabidi MA, Pagani M, Rath M, Schernhuber K, Kazmar T, *et al.*: **Genome-wide assessment of sequence-intrinsic enhancer responsiveness at single-base-pair resolution**. *Nat Biotechnol* 2017, **35**:136–144, https://doi.org/10.1038/nbt.3739.

38. Kheradpour P, Ernst J, Melnikov A, Rogov P, Wang L, Zhang X, *et al.*: **Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay**. *Genome Res* 2013, **23**:800–811, https://doi.org/10.1101/gr.144899.112.

39. Doni Jayavelu N, Jajodia A, Mishra A, Hawkins RD: **Candidate silencer elements for the human and mouse genomes**. *Nat Commun* 2020, **11**:1061, https://doi.org/10.1038/s41467-020-14853-5.

40. * Hofbauer L, Pleyer L-M, Reiter F, Schleiffer A, Vlasova A, Serebreni L, *et al.*: **A genome-wide screen identifies silencers with distinct chromatin properties and mechanisms of repression**. *Mol Cell* 2024, **84**, 45034521.e14, https://doi.org/10.1016/j.molcel.2024.10.041.
This paper developed Silencer-seq, a genome-wide STARR-seq variation that can identify silencer elements in *Drosophila* S2 cells. This highlighted the motifs via which silencers function, and that their chromatin properties make them inaccessible via other common assays.

41. * Mouri K, Dewey HB, Castro R, Berenzy D, Kales S, Tewhey R: **Whole-genome functional characterization of RE1 silencers using a modified massively parallel reporter assay**. *Cell Genomics* 2023, **3**, https://doi.org/10.1016/j.xgen.2022.100234.
The authors developed MPRAduo, an assay to characterize the interaction between two CREs located in *cis* on the same reporter vector to test for repressive element activity.

42. Zhu X, Huang L, Wang C, Li G, Deng B, Kong D, *et al.*: **Uncovering the whole genome silencers of human cells via Ss-STARR-seq**. *Nat Commun* 2025, **16**:723, https://doi.org/10.1038/s41467-025-55852-8.

43. Gisselbrecht SS, Palagi A, Kurland JV, Rogers JM, Ozadam H, Zhan Y, *et al.*: **Transcriptional silencers in Drosophila serve a dual role as transcriptional enhancers in alternate cellular contexts**. *Mol Cell* 2020, **77**:324–337.e8, https://doi.org/10.1016/j.molcel.2019.10.004.

44. * Hong CKY, Wu Y, Erickson AA, Li J, Federico AJ, Cohen BA: **Massively parallel characterization of insulator activity across the genome**. *Nat Commun* 2024, **15**:8350, https://doi.org/10.1038/s41467-024-52599-6.
The authors developed MPIRE, an MPRA technology for recurrently testing CREs at thousands of specific genomic locations simultaneously. MPIRE was able to characterize insulator sequences and determine that their activity is highly context-dependent.

45. Gallego Romero I, Lea AJ: **Leveraging massively parallel reporter assays for evolutionary questions**. *Genome Biol* 2023, **24**:26, https://doi.org/10.1186/s13059-023-02856-6.

46. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, *et al.*: **Systematic localization of common disease-associated variation in regulatory DNA**. *Science* 2012, **337**:1190–1195, https://doi.org/10.1126/science.1222794.

47. Loos RJF: **15 years of genome-wide association studies and no signs of slowing down**. *Nat Commun* 2020, **11**:5900, https://doi.org/10.1038/s41467-020-19653-5.

48. Ulirsch JC, Nandakumar SK, Wang L, Giani FC, Zhang X, Rogov P, *et al.*: **Systematic functional dissection of common genetic variation affecting red blood cell traits**. *Cell* 2016, **165**: 1530–1545, https://doi.org/10.1016/j.cell.2016.04.048.

49. Carrasco Pro S, Hook H, Bray D, Berenzy D, Moyer D, Yin M, *et al.*: **Widespread perturbation of ETS factor binding sites in cancer**. *Nat Commun* 2023, **14**:913, https://doi.org/10.1038/s41467-023-36535-8.

50. Khetan S, Kales S, Kursawe R, Jillette A, Ulirsch JC, Reilly SK, *et al.*: **Functional characterization of T2D-associated SNP effects on baseline and ER stress-responsive β cell transcriptional activation**. *Nat Commun* 2021, **12**:5242, https://doi.org/10.1038/s41467-021-25514-6.

51. Kircher M, Xiong C, Martin B, Schubach M, Inoue F, Bell RJA, *et al.*: **Saturation mutagenesis of twenty disease-associated regulatory elements at single base-pair resolution**. *Nat Commun* 2019, **10**:3583, https://doi.org/10.1038/s41467-019-11526-w.

52. * Lee S, McAfee JC, Lee J, Gomez A, Ledford AT, Clarke D, *et al.*: **Massively parallel reporter assay investigates shared genetic variants of eight psychiatric disorders**. *Cell* 2025, **188**:1409–1424.e21, https://doi.org/10.1016/j.cell.2024.12.022.
This manuscript used MPRAs to test pleiotropic GWAS variants from eight psychiatric disorders, and identified active elements and variants with regulatory activity which were also validated through *in vivo* CRISPR assays in the mouse brain.

53. Lu X, Chen X, Forney C, Donmez O, Miller D, Parameswaran S, *et al.*: **Global discovery of lupus genetic risk variant allelic enhancer activity**. *Nat Commun* 2021, **12**:1611, https://doi.org/10.1038/s41467-021-21854-5.

54. Morova T, Ding Y, Huang C-CF, Sar F, Schwarz T, Giambartolomei C, *et al.*: **Optimized high-throughput screening of non-coding variants identified from genome-wide association studies**. *Nucleic Acids Res* 2023, **51**:e18, https://doi.org/10.1093/nar/gkac1198.

55. Mouri K, Guo MH, de Boer CG, Lissner MM, Harten IA, Newby GA, *et al.*: **Prioritization of autoimmune disease-associated genetic variants that perturb regulatory element activity in T cells**. *Nat Genet* 2022, **54**:603–612, https://doi.org/10.1038/s41588-022-01056-5.

56. Rummel CK, Gagliardi M, Ahmad R, Herholt A, Jimenez-Barron L, Murek V, *et al.*: **Massively parallel functional dissection of schizophrenia-associated noncoding genetic variants**. *Cell* 2023, **186**:5165–5182.e33, https://doi.org/10.1016/j.cell.2023.09.015.

57. Klein JC, Keith A, Agarwal V, Durham T, Shendure J: **Functional characterization of enhancer evolution in the primate lineage**. *Genome Biol* 2018, **19**:99, https://doi.org/10.1186/s13059-018-1473-6.

58. Weiss CV, Harshman L, Inoue F, Fraser HB, Petrov DA, Ahituv N, *et al.*: **The cis-regulatory effects of modern human-specific variants**. *eLife* 2021, **10**, e63713, https://doi.org/10.7554/eLife.63713.

59. * Georgakopoulos-Soares I, Deng C, Agarwal V, Chan CSY, Zhao J, Inoue F, *et al.*: **Transcription factor binding site orientation and order are major drivers of gene regulatory activity**. *Nat Commun* 2023, **14**:2333, https://doi.org/10.1038/s41467-023-37960-5.
The authors employed MPRAs of over 200,000 sequences to investigate TF binding site regulatory grammar and found that orientation and order influence gene regulatory activity.

60. de Boer CG, Vaishnav ED, Sadeh R, Abeyta EL, Friedman N, Regev A: **Deciphering eukaryotic gene-regulatory logic with 100 million random promoters**. *Nat Biotechnol* 2020, **38**:56–65, https://doi.org/10.1038/s41587-019-0315-8.

61. * Sahu B, Hartonen T, Pihlajamaa P, Wei B, Dave K, Zhu F, *et al.*: **Sequence determinants of human gene regulatory elements**. *Nat Genet* 2022, **54**:283–294, https://doi.org/10.1038/s41588-021-01009-4.
The authors used MPRAs to measure transcriptional activity of DNA sequences with 100X human genome coverage and uncovered the rules governing TF grammar through machine learning models.

62. de Almeida BP, Reiter F, Pagani M, Stark A: **DeepSTARR predicts enhancer activity from DNA sequence and enables the de novo design of synthetic enhancers**. *Nat Genet* 2022, **54**: 613–624, https://doi.org/10.1038/s41588-022-01048-5.

63. Pan RW, Röschinger T, Faizi K, Garcia HG, Phillips R: **Deciphering regulatory architectures of bacterial promoters from synthetic expression patterns**. *PLoS Comput Biol* 2024, **20**, e1012697, https://doi.org/10.1371/journal.pcbi.1012697.

64. Barbadilla-Martínez L, Klaassen N, Franceschini-Santos VH, Breda J, Hernandez-Quiles M, Lieshout T van, *et al.*: *The regulatory grammar of human promoters uncovered by MPRA-trained deep learning 2024.* 2024, https://doi.org/10.1101/2024.07.09.602649. 07.09.602649.

65. de Almeida BP, Schaub C, Pagani M, Secchia S, Furlong EEM, Stark A: **Targeted design of synthetic enhancers for selected tissues in the Drosophila embryo**. *Nature* 2024, **626**:207–211, https://doi.org/10.1038/s41586-023-06905-9.

66. ** Gosai SJ, Castro RI, Fuentes N, Butts JC, Mouri K, Alasoadura M, *et al.*: **Machine-guided design of cell-type-targeting cis-regulatory elements**. *Nature* 2024, **634**:1211–1220, https://doi.org/10.1038/s41586-024-08070-z.
This paper trained a deep-learning model to predict synthetic CRE activity in three different cell lines, which helped generate thousands of synthetic CREs with tissue- and cell type-specific characteristics.

67. Zhao J, Baltoumas FA, Konnaris MA, Mouratidis I, Liu Z, Sims J, *et al.*: **MPRAbase a massively parallel reporter assay database**. *Genome Res* 2025, https://doi.org/10.1101/gr.280387.124. gr.280387.124.

68. Yao D, Tycko J, Oh JW, Bounds LR, Gosai SJ, Lataniotis L, *et al.*: **Multicenter integrated analysis of noncoding CRISPRi screens**. *Nat Methods* 2024, **21**:723–734, https://doi.org/10.1038/s41592-024-02216-7.

69. Reilly SK, Gosai SJ, Gutierrez A, Mackay-Smith A, Ulirsch JC, Kanai M, *et al.*: **Direct characterization of cis-regulatory elements and functional dissection of complex genetic associations using HCR-FlowFISH**. *Nat Genet* 2021, **53**:1166–1176, https://doi.org/10.1038/s41588-021-00900-4.

70. ** Zhao S, Hong CKY, Myers CA, Granas DM, White MA, Corbo JC, *et al.*: **A single-cell massively parallel reporter assay detects cell-type-specific gene regulation**. *Nat Genet* 2023, **55**:346–354, https://doi.org/10.1038/s41588-022-01278-7.
The authors developed a single cell MPRA to assess the activity of CREs across cell types and mouse retinal tissue in parallel. Through it, they identified that promoter variants can result in cell-type specific effects on CRE activity.

71. ** Lalanne J-B, Regalado SG, Domcke S, Calderon D, Martin BK, Li X, *et al.*: **Multiplex profiling of developmental cis-regulatory elements with quantitative single-cell expression reporters**. *Nat Methods* 2024, **21**:983–993, https://doi.org/10.1038/s41592-024-02260-3.
In this paper, the authors developed a single-cell assays for MPRA libraries, where a dual RNA reporter successfully separates detection and quantification, providing a highly sensitive multiplexed reporter expression measurement.

72. Kwasnieski JC, Mogno I, Myers CA, Corbo JC, Cohen BA: **Complex effects of nucleotide variants in a mammalian cis-regulatory element**. *Proc Natl Acad Sci* 2012, **109**: 19498–19503, https://doi.org/10.1073/pnas.1210678109.

73. White MA, Myers CA, Corbo JC, Cohen BA: **Massively parallel in vivo enhancer assay reveals that highly local features determine the cis-regulatory function of ChIP-seq peaks**. *Proc Natl Acad Sci U S A* 2013, **110**:11952–11957, https://doi.org/10.1073/pnas.1307449110.

74. Brown AR, Fox GA, Kaplow IM, Lawler AJ, Phan BN, Gadey L, *et al.*: **An in vivo systemic massively parallel platform for deciphering animal tissue-specific regulatory function**. *Front Genet* 2025, **16**, 1533900, https://doi.org/10.3389/fgene.2025.1533900.

75. Chang T-Y, Waxman DJ: **HDI-STARR-seq: condition-specific enhancer discovery in mouse liver in vivo**. *BMC Genom* 2024, **25**:1240, https://doi.org/10.1186/s12864-024-11162-9.

76. McAfee JC, Bell JL, Krupa O, Matoba N, Stein JL, Won H: **Focus on your locus with a massively parallel reporter assay**. *J Neurodev Disord* 2022, **14**:50, https://doi.org/10.1186/s11689-022-09461-x.

77. Johnson GD, Barrera A, McDowell IC, D'Ippolito AM, Majoros WH, Vockley CM, *et al.*: **Human genome-wide measurement of drug-responsive regulatory activity**. *Nat Commun* 2018, **9**:5317, https://doi.org/10.1038/s41467-018-07607-x.

78. Mulvey B, Dougherty JD: **Transcriptional-regulatory convergence across functional MDD risk variants identified by massively parallel reporter assays**. *Transl Psychiatry* 2021, **11**: 1–13, https://doi.org/10.1038/s41398-021-01493-6.

79. * Zahm AM, Owens WS, Himes SR, Fallon BS, Rondem KE, Gormick AN, *et al.*: **A massively parallel reporter assay library to screen short synthetic promoters in mammalian cells**. *Nat Commun* 2024, **15**, 10353, https://doi.org/10.1038/s41467-024-54502-9.

This paper employed a MPRA library of over 6000 synthetic promoters as transcriptional readouts for cellular stimuli to aid systematic development of tunable reporters. The library was validated across cell lines and stimuli and proved particularly efficient in the context of signals mediated by drug targets, such as GPCRs.

80. * Trauernicht M, Filipovska T, Rastogi C, van Steensel B: **Optimized reporters for multiplexed detection of transcription factor activity**. *Cell Syst* 2024, **15**:1107–1122.e7, https://doi.org/10.1016/j.cels.2024.11.003.

The authors employed an MPRA to develop and optimize a library of 86 TFs as transcriptional reporters that can specifically respond to a wide array of perturbations such as stress, metabolism, and immune signalling.