

Identifying Novel Enhancer Elements with CRISPR-Based Screens

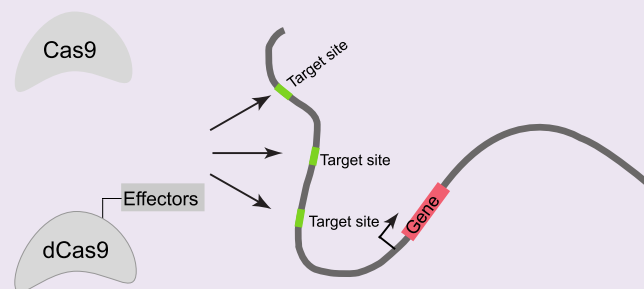
Jason C. Klein,^{*,†} Wei Chen,[‡] Molly Gasperini,[†] and Jay Shendure^{*,†,§}

[†]Department of Genome Sciences, University of Washington, Seattle, Washington 98195, United States

[‡]Molecular Engineering & Science Institute, University of Washington, Seattle, Washington 98195, United States

[§]Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, United States

ABSTRACT: Enhancers control the spatiotemporal expression of genes and are essential for encoding differentiation and development. Since their discovery more than three decades ago, researchers have largely studied enhancers removed from their genomic context. The recent adaptation of CRISPR/Cas9 to genome editing in higher organisms now allows researchers to perturb and test these elements in their genomic context, through both mutation and epigenetic modulation. In this Perspective, we discuss recent advances in scanning noncoding regions of the genome for enhancer activity using CRISPR-based tools.



Although all cells of a multicellular organism share the same genes, the relative timing and levels of expression of these genes somehow specify myriad cell types. The spatiotemporal control of gene expression is influenced by distal DNA sequences known as enhancers. Enhancers were first defined in 1981 as short sequences of DNA that are able to increase the expression of a gene independent of their relative position or orientation to the transcriptional start site.^{1,2} Recently, several groups have shown that disease-associated variants disproportionately reside in candidate enhancer elements.^{3–5}

Despite intensive efforts, identifying and validating enhancers, as well as predicting the effects of sequence variants within them, remain as fundamental challenges for the field. Traditionally, candidate enhancer elements have been characterized by testing their ability to activate expression of a reporter gene (e.g., luciferase or β -galactosidase) in an episomal construct.⁶ Although this type of assay has been considered a gold standard for enhancer function, a major criticism is that it tests sequences outside of their native context.

In recent years, two high-throughput approaches have emerged to identify or validate enhancers. The first approach is to survey genomes for biochemical marks that are associated with enhancer activity (e.g., as done by the ENCODE Consortium on many cell lines). These marks include EP300 ChIP-seq, H3K27ac ChIP-seq, H3K4me1 ChIP-seq, DNase I hypersensitivity (DHS), and others.⁷ While scalable, these assays are fundamentally descriptive and do not measure enhancer activity. The second approach, massively parallel reporter assays (MPRAs), builds on the *in vitro* luciferase assay. However, instead of measuring reporter-gene activity of a single target, MPRAs rely on sequencing-based quantification of thousands of RNA barcodes in parallel, each associated with a different candidate enhancer.^{8–10} MPRAs, and a recent derivative, STARR-seq,¹¹ have been used for saturation mutagenesis of known promoters and enhancers,^{8–10} dissecting

enhancer logic,^{12,13} identifying functional bases and motifs within enhancers,^{14,15} testing the effects of variants on enhancer function,^{16–18} genome annotation,^{11,19} and validating biochemically predicted enhancer elements.^{20,21} While MPRAs directly measure enhancer function and are scalable, they traditionally test short DNA fragments on a plasmid with a minimal promoter, therefore missing the extended sequence context, including its chromatin landscape and any pairing of the candidate enhancer with its endogenous promoter across a long distance. We and colleagues recently described a strategy (“lentiMPRA”) wherein libraries are integrated randomly to the genome,²¹ addressing some but not all of the limitations of MPRAs.

Rather than testing candidate enhancer elements for their positive activity when removed from context, a complementary approach is to perturb these same sequences in their endogenous genomic locations. Until recently, these types of experiments in the laboratory have been limited by the difficulty of engineering targeted perturbations. In 2013, two groups adapted the bacterial CRISPR/Cas9 innate immune system to selectively create targeted double-stranded breaks in mammalian genomes.^{22,23} Due to the ease of cloning the guide sequences, which target the CRISPR complex to specific positions in the genome, Shalem *et al.* applied CRISPR as a genome-wide screen for essential genes in 2014.²⁴ This perspective will focus on a new wave of literature implementing the CRISPR/Cas9 system to perturb and screen thousands to millions of genomic bases for enhancer function in their endogenous sequence contexts.

Special Issue: Chemical Biology of CRISPR

Received: September 6, 2017

Accepted: January 4, 2018

Published: January 4, 2018

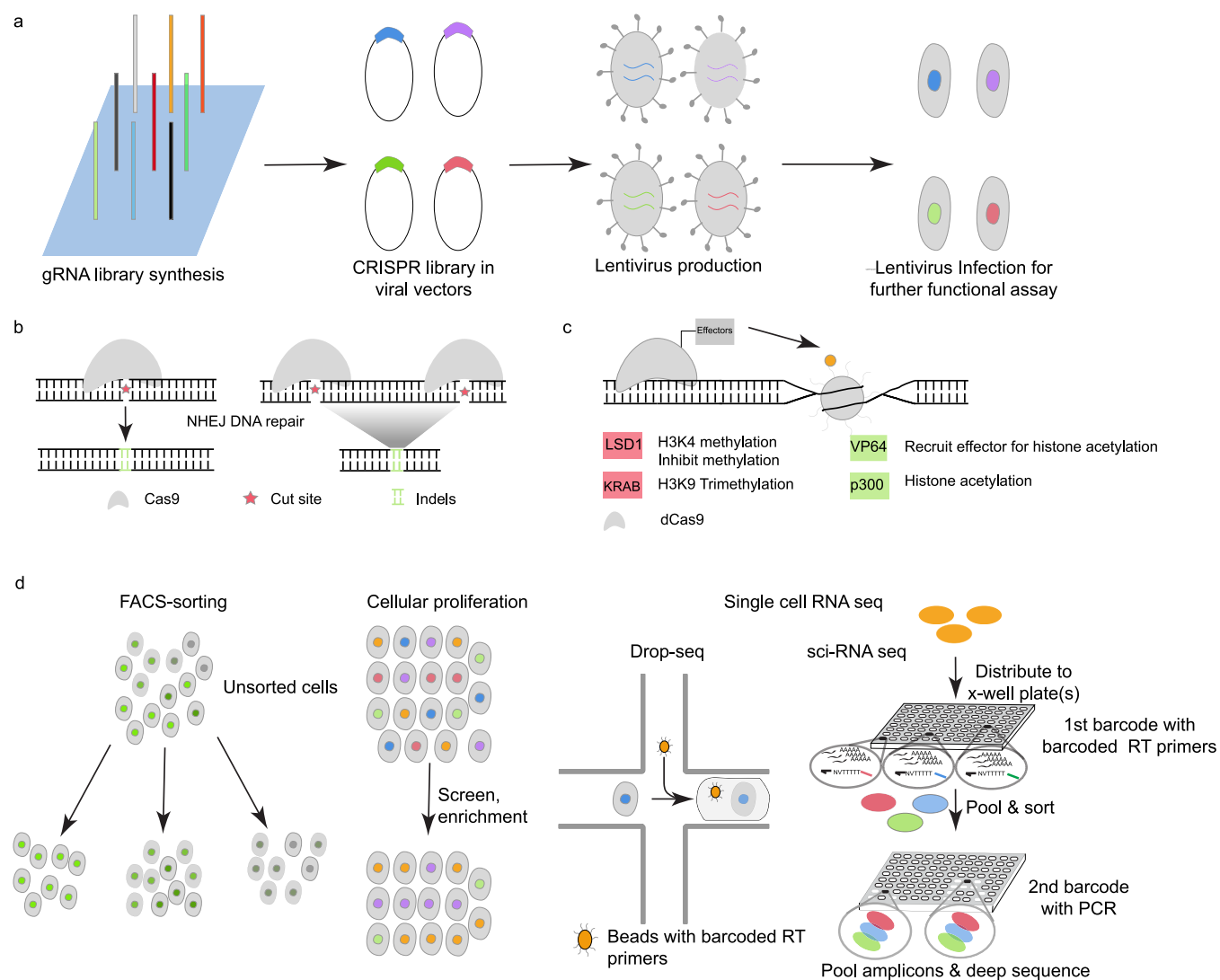


Figure 1. CRISPR-based functional screening for enhancer elements. (a) Designed gRNAs targeting potential enhancer regions are synthesized on a microarray and cloned into viral constructs, which are used to construct lentivirus libraries. Cells are infected with the lentiviral library at a low MOI to confer stable expression of a single construct per cell. (b) Sequence disruption: Putative enhancer regions are either targeted by an individual or paired gRNA construct. Individual gRNAs induce double stranded breaks, which can be repaired by nonhomologous end joining (NHEJ) resulting in short insertions and deletions. Paired gRNAs can result in a drop-out of the intervening sequence, resulting in targeted, large deletions. (c) Epigenetic perturbation: Inactive CRISPR-Cas9 (dCas9) can be fused with effectors for epigenetic modification. LSD1 and KRAB have been used to repress transcription while VP64 and p300 have been used to activate transcription. Effectors directly or indirectly act on histones by introducing histone marks, such as H3K4/H3K9 methylation or H3K27 acetylation. (d) Effects of enhancer perturbation have been measured in at least three ways: Flow sorting based on expression of fluorescently tagged endogenous genes, cellular proliferation advantage/disadvantage controlled by genes of interest, and single cell RNA-seq. Sci-RNA-seq figure adapted from ref 57 with permission from AAAS. Functional enhancers are detected by the identification of the enrichment/depletion of gRNAs or by gene expression changes.

■ TILING DELETION SCANNING

The relatively short guide sequences used to target Cas9 to its desired target sites facilitates large scale knockout screens. Unlike previous genome engineering techniques such as zinc fingers and TALENs, which are laborious to synthesize, tens of thousands of guide RNA (gRNA) sequences can be synthesized in parallel by array-based oligonucleotide synthesis. These gRNAs are then cloned into a library of lentiviral vectors, which each deliver one gRNA into Cas9-expressing cells (Figure 1a).²⁴ Each cell receives its own specific gRNA and resulting programmed mutation; when a functional selection is applied on the diverse pool of cells, the relative abundance of each gRNA, before and after selection, can be readily quantified by sequencing (Figure 1d).

In 2015, Canver *et al.* tiled individual gRNAs across a noncoding region to look for *cis*-regulatory elements.²⁵ A single Cas9 cleavage followed by nonhomologous end joining (NHEJ) results in a spectrum of insertions and deletions (indels) at the target site. These individual gRNA tiling experiments assume that short indels, when occurring at a critical location, will disrupt enhancer function (Figure 1b). Canver *et al.* tiled 3 DHS sites within a previously described *BCL11A* composite enhancer, totaling 3917 nucleotides, with 1130 gRNAs. Since reduction of *BCL11A* results in an increase in HbF, the group sorted cells on the HbF level and examined the prevalence of each gRNA in high versus low HbF populations. The vast majority of gRNAs showed neither enrichment nor depletion, but the approach revealed discrete

genomic loci within DHSs that carry clusters of enriched gRNAs in the HbF high population, indicating that when disrupted, these loci reduce *BCL11A* enhancer activity.

Shortly after, several similar screens were published, utilizing up to 18 000 gRNAs to examine 715 kb of genomic sequence.^{26–29} A nuance of each study is how it enriches for gRNAs impacting enhancer function (Figure 1d). Canver *et al.* sorted cells based on HbF expression and studied an enhancer, that when diminished, increased HbF. Korkmaz *et al.* focused on identifying enhancers regulated by p53 and ER α and selected based on oncogene-induced senescence and ER α expression.²⁹ Sanjana *et al.* focused on identifying enhancers surrounding *NF1*, *NF2*, and *CUL3* and selected based on Vemurafenib resistance.²⁸ All of these assays are limited to only screen for enhancers affecting a particular gene or pathway. Rajagopal *et al.* and Diao *et al.* developed more generalizable strategies by creating fluorescently tagged reporter cell lines and screening for enhancers that modulate expression of the reporter.^{26,27,30}

However, there are several limitations of screening for functional enhancers with short indels created by individual gRNAs. First, the small 1–20 bp deletions introduced through NHEJ^{31,32} may be insufficient to disrupt enhancer function. Both comparisons of enhancer conservation between species³³ and reporter assays with synthetic sequences¹³ have supported a “billboard” model, where at least for some enhancers, a collection of transcription factor binding sites can drive activity in different orders or orientations.³⁴ As such, small deletions may be insufficient to knock down function of some bona fide enhancers. Second, some critical sites may simply not be targeted, as current gRNA design is constrained by the location of PAM sites and other considerations. Finally, because the screens are noisy, these studies aggregate counts from multiple gRNA targets within a sliding window, reducing resolution and power.

Whereas introducing an individual gRNA can create short indels through NHEJ, two groups demonstrated in 2013 that introducing two gRNAs in close proximity can result in a drop-out of the intervening sequence (Figure 1b).^{22,35} By pairing two gRNAs on the same lentiviral vector, libraries of gRNA pairs can create large deletions of programmed size.^{36,37} Zhu *et al.* relied on this for a screen in 2016 that used paired gRNAs to program deletions of 700 human long noncoding RNAs, identifying 51 which can regulate human cancer cell growth.³⁸ In 2017, Diao *et al.* applied this type of screen to search for enhancers around the *Oct4* locus,³⁹ while we and colleagues applied it to scan the *HPRT1* locus.⁴⁰

In 2017, Diao *et al.* tested a library of paired gRNAs on the GFP-tagged *Oct4* cell line published in 2016.³⁹ The study tiled 2 Mb of genomic DNA in human embryonic stem cells with kilobase-sized deletions and identified 45 *cis*-regulatory elements. Gasperini *et al.* scanned a 206 kb region around *HPRT1* with kilobase-sized deletions and selected for loss of HPRT function with 6-thioguanine.⁴⁰ The HPRT screen deviated from the previous designs in two potentially important ways. First, by utilizing overlapping deletions, the screen disrupted each base a median of 27 times, providing additional strength and confidence in calls. Second, in addition to sequencing the gRNAs themselves, Gasperini *et al.* used long-read sequencing to directly sequence editing events as part of postscreen validation. The study found that *HPRT1* was largely robust to noncoding deletions, concluding that the proximal regulatory sequence was sufficient for *HPRT1* expression and

direct sequencing of selected edits is important for reducing false positives in CRISPR-based screens of noncoding sequence.

■ TILING EPIGENETIC MODIFICATION SCANNING

A highly related approach to deletion scanning is to instead use CRISPR to modify the epigenetic landscape around candidate enhancer sequences. Since Cas9 can be targeted to almost any region of the genome, Qi *et al.* developed a catalytically inactive version of Cas9 (dCas9), which can function as an RNA-guided DNA recognition platform.⁴¹ Several groups have since fused dCas9 to repressor and activator domains to modulate expression of target genes (Figure 1c). Here, we only focus on the effectors that have been used to target enhancers, including two repressors (KRAB, LSD1) and two activators (VP64 and p300).

The Krüppel-associated box (KRAB) domain is the most commonly used repressor for dCas9 experiments. KRAB recruits cofactors that repress transcription through histone methylation and deacetylation.^{42–46} Gilbert *et al.* targeted dCas9-KRAB fusions to promoters in order to repress gene expression.⁴⁷ Since then, several groups have targeted dCas9-KRAB to putative enhancers in order to validate regulatory function in the genome. These studies used dCas9-KRAB to target previously described enhancers of *Nanog*,⁴⁸ *Oct4*,^{48,49} *Tbx3*,⁴⁹ and hemoglobin subunit genes.⁴² Kearns *et al.* compared the roles of KRAB and lysine-specific histone demethylase 1 (LSD1) when fused to *Neisseria meningitidis* dCas9 to target regulatory sequences.⁴⁹ LSD1 is a chromatin regulator that has been proposed to silence enhancers during embryonic stem cell differentiation by demethylating histone H3 on lysine 4 or lysine 9.⁵⁰ While dCas9-KRAB repressed expression when targeted to promoters, proximal enhancers, and distal enhancers, dCas9-LSD1 only repressed gene expression when targeted to distal enhancers. Ultimately, utilizing both of these orthogonal proteins and more may add additional sensitivity and specificity to future screens.

While the previous papers targeted known enhancers, Fulco *et al.* used dCas9-KRAB to scan 1.29 Mb of genomic sequence with 98 000 gRNAs around *GATA1* and *MYC* in K562 erythroleukemia cells.⁵¹ Since *GATA1* and *MYC* affect proliferation of these cells, gRNAs that disrupt enhancer function would be depleted after cell proliferation. Through this screen, they identified two and seven distal regulatory elements for *GATA1* and *MYC*, respectively. Similar to many of the deletion scans, however, this functional assay is limited to identifying regulators of genes important in proliferation and therefore not generalizable to all loci.

In an attempt to create a more generalizable assay, Klann *et al.* labeled their genes of interest with an mCherry-tagged reporter, similar to Rajagopal *et al.* and Diao *et al.*'s use of GFP-tagging in individual gRNA scanning. While a fluorescent tag allows selection on a gene without a known selectable function, it is not scalable for annotating several genes, as a new cell line has to be created for each gene of interest in each cell line of interest. For example, to study all genes in five cell lines, one would need over 100 000 genetically engineered cell lines.

Xie *et al.* attempted to create a more scalable, generalizable assay by utilizing single cell RNA-sequencing to phenotype enhancer perturbations. Mosaic single-cell analysis by indexed CRISPR sequencing (MOSAIC-seq) uses single cell RNA-sequencing as a global measurement of differential gene expression.⁵² The group designed 241 gRNAs targeting

Table 1

	individual guide deletion scanning	paired guide deletion scanning	epigenetic modification scanning
mechanism of action	<ul style="list-style-type: none"> • Cas9 cleavage followed by NHEJ-mediated insertions and deletions • short ~1–20 bp insertions/deletions 	<ul style="list-style-type: none"> • Cas9 cleavage followed by NHEJ-mediated drop out of intervening sequence • large deletions of programmed size (kilobase-sized) 	<ul style="list-style-type: none"> • repression: histone methylation and deacetylation (KRAB) or histone demethylation (LSD1) • activation: introducing scaffold for preinitiation complex (VP64) or direct acetylation (P300)
potential sources of false positives	<ul style="list-style-type: none"> • rare deletions may unexpectedly extend into nearby coding sequence, promoter, or UTR • deletions may introduce novel TF binding sites 	<ul style="list-style-type: none"> • rare, unprogrammed deletions may unexpectedly extend into nearby coding sequence, promoter, or UTR 	<ul style="list-style-type: none"> • synthetic effector could unexpectedly spread to nearby enhancer or in 3D space to promoter^{42,55}
potential sources of false negatives	<ul style="list-style-type: none"> • small deletions may not disrupt robust enhancers • redundancy/compensation by nearby, unperturbed enhancer 	<ul style="list-style-type: none"> • lower editing rate compared to individual guide screens • redundancy/compensation by nearby, unperturbed enhancer 	<ul style="list-style-type: none"> • redundancy/compensation by nearby, unperturbed enhancer • effectors may not be able to disrupt activity of certain enhancers
references	25–29	38–40	30, 42, 48, 49, 51–54

dCas9-KRAB to 71 constituent enhancers from 15 superenhancers—large regions composed of multiple predicted enhancers. Surprisingly, only one to two enhancers within each superenhancer significantly reduced target gene expression. The enhancers that decreased target gene expression were significantly enriched for RNA polymerase II and p300 binding and only showed moderate enrichment for enhancer-associated histone modifications, H3K4me1 and H3K27ac. While MOSAIC-seq is much more generic than other assays, the detection limit and cost of single-cell sequencing may reduce its utility in practice to genes that are highly expressed.

Activators have also been fused to dCas9 in order to increase target gene expression. Similar to KRAB, dCas9-VP64 was first used to target promoter regions. Both Gao *et al.* and Hilton *et al.* targeted distal enhancers with dCas9-VP64 and showed moderate gene activation.^{48,53} Simeonov *et al.* recently screened for novel enhancers of *CD69* and *IL2RA* using dCas9-VP64. The group tiled 135kb around *CD69* with over 10 000 gRNAs and 178kb around *IL2RA* with over 20 000 gRNAs and identified several dCas9-VP64-responsive elements.⁵⁴ Hilton *et al.* also fused p300, the catalytic core of human acetyltransferase, to dCas9 (dCas9-p300^{Core}) to target enhancers.⁵³ dCas9-p300^{Core} significantly enhanced target gene expression, potentially by means of acetylation on histone H3 on lysine 27 (H3K27ac). dCas9-p300^{Core} was highly specific and able to induce robust activation with only one guide RNA, making it a candidate tool for genome-wide screening. Recently, Klann *et al.* developed CRISPR-Cas9-based epigenomic regulatory element screening (CERES), which combines dCas9-p300^{Core} and dCas9-KRAB to obtain both gain and loss of function information by targeting the same regions with a repressor and an activator.³⁰ They targeted a 4-Mb region including 433 DHSs surrounding *HER2* with a library of 12 189 gRNAs and measured *HER2* expression using immunofluorescence staining. Loss and gain of function assays were performed in two different cell lines—A431 epidermoid carcinoma cells with moderate *HER2* expression and HEK293T cells with low *HER2* expression, respectively. With the same gRNA library, results from A431 and HEK293T cells were highly correlated with mirrored effects, same trend but opposite direction, providing additional confidence, which is critical for regions with small effects on transcription.

FUTURE DIRECTIONS

Functional characterization of enhancer elements is imperative for advancing our understanding of how and when genes are

expressed, which in turn is needed to understand how cells differentiate, how species evolve, and how noncoding variants contribute to phenotype and disease. While biochemical annotations and MPRAs have allowed researchers to begin answering these questions, they have done so in a context far-removed from the native genome. Finally, with the advent of CRISPR screens, we can begin to overcome this limitation.

Even so, we are not quite ready to apply these screens to entire genomes or to multiple cell types, which each provide a unique *trans* environment for enhancers to act. First, we need a better understanding of the relative sensitivities and specificities of the different CRISPR-based screens. We highlight some potential limitations of each assay in Table 1. However, for a true comparison, we need studies that screen the same locus or loci with multiple assays. Although it will be a valuable data set for this nascent field moving forward, such a systematic comparison has yet to be conducted or published. Moreover, in order to reach these goals, noncoding CRISPR screens must become more generalizable, higher throughput, and less expensive. The GFP and mCherry-tagging screens described above could in theory be applied to every gene but would require the generation of thousands of different engineered cell lines followed by thousands of independent experiments. Coupling CRISPR-based perturbations to single-cell RNA sequencing, as performed in MOSAIC-seq, provides a widely applicable assay but has its own limitations. Due to the relatively low depth of current single cell RNA sequencing protocols, assaying genes that are not highly expressed would require deep sequencing of many cells. Methods such as Drop-seq and single cell combinatorial indexing RNA-seq (sci-RNA-seq) are making single cell RNA readouts more scalable.^{56,57} As these techniques reduce cost and increase scalability, methods like MOSAIC-seq will become more feasible to perform genome-wide or on multiple cell types.

We emphasize that CRISPR screens (which try to knock out or modulate native sequences, in context) and MPRAs (which test large numbers of sequences for positive activity, independent of context) are complementary rather than competing. While CRISPR-based screens have the advantage of detecting enhancers in their endogenous locations, they lack the resolution needed to routinely screen large numbers of sequences or sequence variants for their isolated/independent effects on expression. CRISPR screens do not remove sequences being tested from their broader genomic context, which is either a disadvantage or an advantage, depending on the question that one is asking. There are likely currently

underexplored opportunities to synergize MPRA and CRISPR-based screens, e.g., to quantify the effects of variants within CRISPR-verified enhancers, or to understand the extent to which enhancer effects are dependent or independent of the broader sequence context. As such, MPRA methods remain indispensable for testing synthetic sequences and sequence variants, which either do not occur in the genome or would be difficult to introduce in high-throughput by genome engineering. For example, MPRA-based methods have been, and will continue to be, critical in studying enhancer logic through saturation mutagenesis^{8–10} and motif shuffling.¹³ While not testing sequences in their endogenous location, our recently published lentiMPRA nonetheless tests them in chromatinized contexts and correlates with biochemical predictions better than traditional plasmid-based MPRA. The further development and application of MPRA as well as CRISPR-based screens will be necessary for our continued progress toward understanding enhancer logic and function.

The scalability of CRISPR has opened the door for novel screens and assays to dissect the noncoding genome. Together with MPRA, we predict that CRISPR screens will play a critical role in elucidating how and when genes are expressed, which remains one of the most important and difficult questions in genomics. In the near term, these assays will improve our ability to annotate the genome for regulatory elements, a critical challenge in the wake of ENCODE as we are currently awash in unverified noncoding regulatory elements based on biochemical marks. The validation of these annotations (or their failure to validate), ideally through some combination of CRISPR and MPRA-based assays, is a critical next step for the field of regulatory genomics. Of note, CRISPR based screens in cell lines, although “in context” in terms of genomic sequence, remain “out of context” in terms of endogenous development. *In vivo* studies—also enabled by CRISPR but still low-throughput—will remain a critical tool as well.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: shendure@u.washington.edu.

*E-mail: jcklein@u.washington.edu.

ORCID

Jason C. Klein: 0000-0001-9566-6347

Funding

This work was funded by grants from the National Institutes of Health (UM1HG009408) to J.S. J.K. was supported in part by 1F30HG009479 from the NHGRI. M.G. was supported by grants from the National Science Foundation (Graduate Research Fellowship) and NIH (5T32HG000035). J.S. is an investigator of the Howard Hughes Medical Institute.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank the Shendure lab, and in particular E. Aguilar and A. Keith for discussions and feedback.

KEYWORDS

Enhancer: Classically defined as short sequences of DNA, which are able to increase expression of a gene independent of their relative position or orientation to the transcriptional start site

CRISPR: Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) is a bacterial defense system that recognizes and destroys foreign DNA. CRISPR has recently been modified for genome engineering in mammalian cells.

Epigenetics: Usually biochemical modifications to DNA or relevant proteins that cause heritable changes in gene function without changing the nucleotide sequence. In the context of this review, we are mainly focusing on activating and repressing histone modifications.

Massively Parallel Reporter Assay: A plasmid-based assay to measure the regulatory effects on gene expression of thousands of independent sequences at the same time

Guide RNA: An RNA sequence that contains a scaffold for Cas-binding and a spacer sequence that targets DNA. In the context of this review, guide RNAs target Cas9 to specific genomic sites.

Episomal: DNA that replicates independently of chromosomal DNA. While traditional MPRA are episomal assays, CRISPR-based screens identify enhancer elements directly in their endogenous genomic loci.

Nonhomologous end joining: A repair mechanism for double stranded breaks, which directly ligates broken DNA without a template. In the context of this review, the field relies on imperfect NHEJ to create insertions and deletions while repairing double-stranded breaks induced by Cas9.

ChIP-seq: Chromatin immunoprecipitation followed by sequencing of cross-linked DNA is used to identify protein–DNA interactions. In the context of this review, ChIP-seq has been used to identify regions of DNA bound by proteins and modified-histones associated with enhancer activity, such as P300, H3K27ac, and H3K4me1.

REFERENCES

- (1) Banerji, J., Rusconi, S., and Schaffner, W. (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299–308.
- (2) Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M. P., and Chambon, P. (1981) The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res.* 9, 6047–6068.
- (3) Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., Reynolds, A. P., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutayavin, T., Stehling-Sun, S., Johnson, A. K., Canfield, T. K., Giste, E., Diegel, M., Bates, D., Hansen, R. S., Neph, S., Sabo, P. J., Heimfeld, S., Raubitschek, A., Ziegler, S., Cotsapas, C., Sotoodehnia, N., Glass, I., Sunyaev, S. R., Kaul, R., and Stamatoyanopoulos, J. A. (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190–1195.
- (4) Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S., and Snyder, M. (2012) Linking disease associations with regulatory information in the human genome. *Genome Res.* 22, 1748–1759.
- (5) Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., Amin, V., Whitaker, J. W., Schultze, M. D., Ward, L. D., Sarkar, A., Quon, G., Sandstrom, R. S., Eaton, M. L., Wu, Y., Pfening, A. R., Wang, X., Claussnitzer, M., Liu, Y., Coarfa, C., Harris, R. A., Shores, N., Epstein, C. B., Gjoneska, E., Leung, D., Xie, W., Hawkins, R. D., Lister, R., Hong, C., Gascard, P., and Mungall, A. J. (2015) Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317–330.
- (6) Arnone, M. I., Dmochowski, I. J., and Gache, C. (2004) Using reporter genes to study cis-regulatory elements. *Methods Cell Biol.* 74, 621–652.
- (7) Shlyueva, D., Stampfel, G., and Stark, A. (2014) Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* 15, 272–286.

- (8) Patwardhan, R. P., Lee, C., Litvin, O., Young, D. L., Pe'er, D., and Shendure, J. (2009) High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat. Biotechnol.* 27, 1173–1175.
- (9) Patwardhan, R. P., Hiatt, J. B., Witten, D. M., Kim, M. J., Smith, R. P., May, D., Lee, C., Andrie, J. M., Lee, S.-I., Cooper, G. M., Ahituv, N., Pennacchio, L. A., and Shendure, J. (2012) Massively parallel functional dissection of mammalian enhancers in vivo. *Nat. Biotechnol.* 30, 265–270.
- (10) Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C. G., Kinney, J. B., Kellis, M., Lander, E. S., and Mikkelsen, T. S. (2012) Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* 30, 271–277.
- (11) Arnold, C. D., Gerlach, D., Stelzer, C., Boryn, L. M., Rath, M., and Stark, A. (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074–1077.
- (12) Nguyen, T. A., Jones, R. D., Snively, A. R., Pfening, A. R., Kirchner, R., Hemberg, M., and Gray, J. M. (2016) High-throughput functional comparison of promoter and enhancer activities. *Genome Res.* 26, 1023–1033.
- (13) Smith, R. P., Taher, L., Patwardhan, R. P., Kim, M. J., Inoue, F., Shendure, J., Ovcharenko, I., and Ahituv, N. (2013) Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. *Nat. Genet.* 45, 1021–1028.
- (14) Kheradpour, P., Ernst, J., Melnikov, A., Rogov, P., Wang, L., Alston, J., Mikkelsen, T. S., Kellis, M., and Zhang, K. (2013) Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay. *Genome Res.* 23, 800–811.
- (15) Ernst, J., Melnikov, A., Zhang, X., Wang, L., Rogov, P., Mikkelsen, T. S., and Kellis, M. (2016) Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions. *Nat. Biotechnol.* 34, 1180–1190.
- (16) Vockley, C. M., Guo, C., Majoros, W. H., Nodzenski, M., Scholtens, D. M., Hayes, M. G., Lowe, W. L., and Reddy, T. E. (2015) Massively parallel quantification of the regulatory effects of noncoding genetic variation in a human cohort. *Genome Res.* 25, 1206–1214.
- (17) Ulirsch, J. C., Nandakumar, S. K., Wang, L., Giani, F. C., Rogov, P., Melnikov, A., McDonel, P., Do, R., Mikkelsen, T. S., Zhang, X., and Sankaran, V. G. (2016) Systematic functional dissection of common genetic variation affecting red blood cell traits. *Cell* 165, 1530–1545.
- (18) Tewhey, R., Kotliar, D., Park, D. S., Liu, B., Winnicki, S., Steven, K., et al. (2016) Direct identification of hundreds of expression-modulating variants using a multiplexed reporter assay. *Cell* 165, 1519–1529.
- (19) Arnold, C. D., Gerlach, D., Spies, D., Matts, J. A., Sytnikova, Y. A., Pagani, M., Lau, N. C., and Stark, A. (2014) Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat. Genet.* 46, 685–692.
- (20) Kwasniewski, J. C., Fiore, C., Chaudhari, H. G., and Cohen, B. A. (2014) High-throughput functional testing of ENCODE segmentation predictions. *Genome Res.* 24, 1595–1602.
- (21) Inoue, F., Kircher, M., Martin, B., Cooper, G. M., Witten, D. M., McManus, M. T., Ahituv, N., and Shendure, J. (2017) A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome Res.* 27, 38–52.
- (22) Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., Habib, N., and Zhang, F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823.
- (23) Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., Dicarlo, J. E., Norville, J. E., and Church, G. M. (2013) RNA-guided human genome engineering via Cas9. *Science* 339, 823–826.
- (24) Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., and Zhang, F. (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84–87.
- (25) Canver, M. C., Smith, E. C., Sher, F., Pinello, L., Sanjana, N. E., Shalem, O., Chen, D. D., Schupp, P. G., Vinjamur, D. S., Garcia, S. P., Luc, S., Kurita, R., Nakamura, Y., Fujiwara, Y., et al. (2015) BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527, 192–197.
- (26) Rajagopal, N., Srinivasan, S., Kooshesh, K., Guo, Y., Edwards, M. D., Banerjee, B., Syed, T., Emons, B. J. M., Gifford, D. K., and Sherwood, R. I. (2016) High-throughput mapping of regulatory DNA. *Nat. Biotechnol.* 34, 167–174.
- (27) Diao, Y., Li, B., Meng, Z., Jung, I., Lee, A. Y., Dixon, J., Maliskova, L., Guan, K., Shen, Y., and Ren, B. (2016) A new class of temporarily phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening. *Genome Res.* 26, 397–405.
- (28) Sanjana, N. E., Wright, J., Zheng, K., Shalem, O., Fontanillas, P., Joung, J., Cheng, C., Regev, A., and Zhang, F. (2016) High-resolution interrogation of functional elements in the noncoding genome. *Science* 353, 1545–1549.
- (29) Korkmaz, G., Lopes, R., Ugalde, A. P., Nevedomskaya, E., Han, R., Myacheva, K., Zwart, W., Elkon, R., and Agami, R. (2016) Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. *Nat. Biotechnol.* 34, 192–198.
- (30) Klann, T. S., Black, J. B., Chellappan, M., Safi, A., Song, L., Hilton, I. B., Crawford, G. E., Reddy, T. E., and Gersbach, C. A. (2017) CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat. Biotechnol.* 35, 561–568.
- (31) Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A. J., Le, L. P., Aryee, M. J., and Joung, J. K. (2014) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187–197.
- (32) McKenna, A., Findlay, G. M., Gagnon, J. A., Horwitz, M. S., Schier, A. F., and Shendure, J. (2016) Whole organism lineage tracing by combinatorial and cumulative genome editing. *Science* 353, 462–473.
- (33) Taher, L., Mcgaughey, D. M., Maragh, S., Aneas, I., Bessling, S. L., Miller, W., Nobrega, M. A., McCallion, A. S., and Ovcharenko, I. (2011) Genome-wide identification of conserved regulatory function in diverged sequences. *Genome Res.* 21, 1139–1149.
- (34) Arnosti, D. N., and Kulkarni, M. M. (2005) Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *J. Cell. Biochem.* 94, 890–898.
- (35) Yang, H., Wang, H., Shivalila, C. S., Cheng, A. W., Shi, L., and Jaenisch, R. (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370–1379.
- (36) Vidigal, J. A., and Ventura, A. (2015) Rapid and efficient one-step generation of paired gRNA CRISPR-Cas9 libraries. *Nat. Commun.* 6, 8083.
- (37) Aparicio-Prat, E., Arnan, C., Sala, I., Bosch, N., Guigó, R., and Johnson, R. (2015) DECKO: Single-oligo, dual-CRISPR deletion of genomic elements including long non-coding RNAs *BMC Genomics* 16, DOI: 10.1186/s12864-015-2086-z.
- (38) Zhu, S., Li, W., Liu, J., Chen, C., Liao, Q., Xu, P., Xu, H., Xiao, T., et al. (2016) Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. *Nat. Biotechnol.* 34, 1279–1286.
- (39) Diao, Y., Fang, R., Li, B., Meng, Z., Yu, J., Qiu, Y., Lin, K. C., Huang, H., Liu, T., Marina, R. J., Jung, I., Shen, Y., Guan, K., and Ren, B. (2017) A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat. Methods* 14, 629–635.
- (40) Gasperini, M., Findlay, G. M., McKenna, A., Milbank, J. H., Lee, C., Zhang, M. D., Cusanovich, D. A., and Shendure, J. (2017) CRISPR/Cas9-mediated scanning for regulatory elements required for HPRT1 expression via thousands of large, programmed genomic deletions. *Am. J. Hum. Genet.* 101, 192–205.
- (41) Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013) Repurposing CRISPR as an

RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183.

(42) Thakore, P. I., D'Ippolito, A. M., Song, L., Safi, A., Shivakumar, N. K., Kabadi, A. M., Reddy, T. E., Crawford, G. E., and Gersbach, C. A. (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* 12, 1143–1149.

(43) Sripathy, S. P., Stevens, J., and Schultz, D. C. (2006) The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zing finger protein-mediated transcriptional repression. *Mol. Cell. Biol.* 26, 8623–8638.

(44) Groner, A. C., Meylan, S., Ciuffi, A., Zangger, N., and Ambrosini, G. (2010) KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet.* 6, e1000869.

(45) Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G., and Rauscher, F. J., 3rd (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* 16, 919–932.

(46) Reynolds, N., Dvinge, H., Hynes-allen, A., Balasooriya, G., Leaford, D., Behrens, A., Bertone, P., Hendrich, B., and Salmon-Divon, M. (2012) NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. *EMBO J.* 31, 593–605.

(47) Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., and Qi, L. S. (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451.

(48) Gao, X., Tsang, J. C. H., Gaba, F., Wu, D., Lu, L., and Liu, P. (2014) Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Res.* 42, e155.

(49) Kearns, N. A., Pham, H., Tabak, B., Genga, R. M., Silverstein, N. J., Garber, M., and Maehr, R. (2015) Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* 12, 401–403.

(50) Whyte, W. A., Bilodeau, S., Orlando, D. A., Hoke, H. A., Frampton, G. M., Foster, C. T., Cowley, S. M., and Young, R. A. (2012) Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* 482, 221–225.

(51) Fulco, C. P., Munschauer, M., Anyoha, R., Munson, G., Grossman, S. R., Perez, E. M., Kane, M., Cleary, B., Lander, E. S., and Engreitz, J. M. (2016) Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* 354, 769–773.

(52) Xie, S., Duan, J., Li, B., Zhou, P., and Hon, G. C. (2017) Multiplexed engineering and analysis of combinatorial enhancer activity in single cells. *Mol. Cell* 66, 285–299.

(53) Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., and Gersbach, C. A. (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33, 510–517.

(54) Simeonov, D. R., Gowen, B. G., Boontanrart, M., Roth, T. L., Gagnon, J. D., Mumbach, M. R., Satpathy, A. T., Lee, Y., Bray, N. L., Chan, A. Y., Lituiev, D. S., Nguyen, M. L., Gate, R. E., Subramaniam, M., Li, Z., Woo, J. M., Mitros, T., Huang, H., Liu, R., Tobin, V. R., Schumann, K., Daly, M. J., Farh, K. K., Ansel, K. M., Ye, C. J., Greenleaf, W. J., Anderson, M. S., Bluestone, J. A., Chang, H. Y., Corn, J. E., Marson, A., et al. (2017) Discovery of stimulation-responsive enhancers with CRISPR activation. *Nature* 549, 111–115.

(55) Lopes, R., Korkmaz, G., and Agami, R. (2016) Applying CRISPR-Cas9 tools to identify and characterize transcriptional enhancers. *Nat. Rev. Mol. Cell Biol.* 17, 597–604.

(56) Macosko, E. Z., Basu, A., Satija, R., Nemeshe, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A., and

McCarroll, S. A. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161, 1202–1214.

(57) Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S. N., Steemers, F. J., Adey, A., Waterston, R. H., Trapnell, C., and Shendure, J. (2017) Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 357, 661–667.