The Company of Biologists

REVIEW

Using CRISPR to understand and manipulate gene regulation

Ersin Akinci[1,2,*], Marisa C. Hamilton[1,*], Benyapa Khowpinitchai[1] and Richard I. Sherwood[1,3,‡]

ABSTRACT

Understanding how genes are expressed in the correct cell types and at the correct level is a key goal of developmental biology research. Gene regulation has traditionally been approached largely through observational methods, whereas perturbational approaches have lacked precision. CRISPR-Cas9 has begun to transform the study of gene regulation, allowing for precise manipulation of genomic sequences, epigenetic functionalization and gene expression. CRISPR-Cas9 technology has already led to the discovery of new paradigms in gene regulation and, as new CRISPR-based tools and methods continue to be developed, promises to transform our knowledge of the gene regulatory code and our ability to manipulate cell fate. Here, we discuss the current and future application of the emerging CRISPR toolbox toward predicting gene regulatory network behavior, improving stem cell disease modeling, dissecting the epigenetic code, reprogramming cell fate and treating diseases of gene dysregulation.

KEY WORDS: CRISPR screening, CRISPR-Cas9, Disease modeling, Epigenetics, Gene regulation

Introduction

Every organism is endowed with a distinct genome sequence that holds the instructions to build a body plan consisting of thousands of different cell types capable of carrying out distinct functions and responding to changing external conditions. Understanding how the genome orchestrates cell type- and condition-specific expression of genes is one of the major challenges of developmental biology.

Deciphering gene regulation has lagged behind deciphering gene function because it is an inherently more difficult problem. Gene function relies on a simple genetic code of 64 codons that was cracked in the 1960s (Nirenberg et al., 1966) and a splicing code based on singular splice donor and acceptor consensus motifs that, although more complex than the genetic code, is also largely understood (Barash et al., 2010). It is now possible to predict with decent accuracy which coding and splicing mutations will or will not disrupt gene function (Jaganathan et al., 2019; Kircher et al., 2014). Gene regulation, meanwhile, has not been as easy to solve because it relies on more diverse distributed coding principles. There are over 1600 human transcription factors (TFs), each of which recognizes a set of sequences (motifs; see Glossary, Box 1) based on thermodynamic features of protein-DNA binding affinity (Lambert et al., 2018). This property makes the TF-binding code orders of magnitude more complex than the genetic and splicing codes. Moreover, individual TF motifs are rarely – if ever –

sufficient to enable binding in the genome owing to competition with histones, which hinder accessibility of TFs to DNA; rather, TFs bind combinatorially, using principles that have not been well elucidated (Reiter et al., 2017). Compounding this issue of combinatorics, most TFs have cell-type-specific expression, so valid cofactor combinations vary substantially depending on context. The binding of a TF to an individual DNA regulatory region, in turn, tends to induce only modest impacts on gene expression. Here, again, combinatorics among gene regulatory regions (see Glossary, Box 1) is required to fully explain the expression of any given gene. Finally, gene regulatory interactions are only loosely spatially limited within megabase-scale topologically-associating domains (TADs; see Glossary, Box 1) (Gonzalez-Sandoval and Gasser, 2016), making it difficult to determine *a priori* to which gene(s) a regulatory element belongs. Altogether, the cognate prediction task that is relatively well solved for gene function and splicing – how a given non-coding motif or mutation will alter gene regulation – remains unsolved.

Improving our understanding of gene regulation would have a major impact on our ability to understand, and control, cellular identity and function. Stem cell differentiation relies on recapitulating the gene regulatory process that occurs during development, and improved understanding of gene regulation should lead to improved ability to manipulate stem cell differentiation and reprogramming (Cherry and Daley, 2012; Cohen and Melton, 2011; Qian et al., 2020; Xu et al., 2015; Yeo et al., 2020; Zhou et al., 2020). Beyond cell fate engineering, non-coding genomic sequence variants, which are largely thought to act through gene regulation, underlie the bulk of genetic risk of disease (Khera et al., 2018; Nishizaki and Boyle, 2017). Again, understanding gene regulatory principles would help us predict, and possibly correct, disease-inducing non-coding sequence changes.

In this Review, we advance the view that the emergence of CRISPR-Cas9 (see Glossary, Box 1) genome editing has the potential to accelerate progress towards deciphering gene regulation. Gene regulation research has traditionally been chiefly observational, with large efforts to map transcriptional differences among cell types (He et al., 2020; Regev et al., 2017; Wilbrey-Clark et al., 2020), chromatin states that correlate with gene regulatory status (Roadmap Epigenomics Consortium et al., 2015; Stunnenberg et al., 2016) and TF-binding sites (The ENCODE Project Consortium, 2012). Previous perturbational approaches have either lacked precision, such as studies of animal-wide TF knockouts, or have relied on empirically derived sets of sequence variants, such as analysis of quantitative trait loci (Aguet et al., 2017; Degner et al., 2012; McVicker et al., 2013) that are biased by evolutionary pressures, limited in their ability to dissect the function of sequences in linkage disequilibrium and lacking the scale to examine the full spectrum of possible sequence changes. CRISPR-Cas9 has begun to transform the study of gene regulation, allowing for precise, unbiased manipulation of gene regulatory sequences (see Glossary, Box 1) and the alteration of the expression of any

[1]Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. [2]Department of Agricultural Biotechnology, Faculty of Agriculture, Akdeniz University, Antalya, 07070, Turkey. [3]Hubrecht Institute, 3584 CT, Utrecht, The Netherlands.
*These authors contributed equally to this work

‡Author for correspondence (rsherwood@bwh.harvard.edu)

R.I.S., 0000-0001-7427-7713

DEVELOPMENT

1

Box 1. Glossary

Base editing: CRISPR-based technique in which Cas9-nickase is attached to a deaminase enzyme (cytosine deaminases or adenosine deaminases), which catalyzes the conversion of targeted C:G or A:T base pairs into T:A or G:C, respectively.

CRISPR-Cas9: Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) is a genome editing tool, adapted from naturally occurring bacterial defense system against viruses, and directed to its target DNA region with a single guide RNA sequence (sgRNA) that is able to bind Cas9 nuclease and target DNA.

CRISPRa/i: Gene activating/inhibiting CRISPR tools in which transcriptional activators/inhibitors are attached to dCas9.

dCas9: 'Deactivated' or 'dead' Cas9, which is mutated in the catalytic domain so that it binds but does not cleave the target DNA.

Enhancer: A regulatory DNA region, which may be located distally to target genes, that activates gene expression by binding transcriptional activators.

Epistatic relationship: Interaction between two or more different components, in which the combined function is distinct from the additive function of each individual component.

eRNA: Enhancer RNA, a short non-coding RNA transcribed from an enhancer region.

Gene regulatory regions/elements: Usually non-coding DNA segments such as enhancers, silencers, promoters, topological elements (see below) or insulators that regulate gene expression through interaction with transcription factors and non-coding RNAs.

Genome-wide CRISPR-Cas9 knockout screens: An experimental approach in which sgRNAs are designed to target every gene in the genome, facilitating discovery of key genes or genetic sequences that elicit a specific function or phenotype through Cas9-induced genetic loss-of-function.

GRN: Gene regulatory network comprises a set of interacting trans-regulatory elements such as transcription factors and cis-regulatory elements such as enhancers and promoters that control a biological process such as differentiation.

GWAS: Genome-wide association studies, which use matched genotype and phenotype information from large cohorts of human patients to associate common genetic variants with physiological and disease phenotypes.

Motif: A short sequence of DNA that recurs within the genome and has a biological function, such as binding a transcription factor.

PAM: The protospacer adjacent motif is a short DNA motif in the genome following the sgRNA-targeted sequence, which is necessary for specific recognition of the target genome by Cas9.

PIC-seq: A single-cell sequencing strategy to elucidate the sequence information of physically interacting cells.

Prime editing: A recent CRISPR-based genome editing platform, in which Cas9-nickase is attached to a reverse transcriptase. When paired with a prime editing guide RNA that is extended to include a primer and template for the reverse transcriptase, prime editing is capable of inducing all possible base-to-base conversions and short indels without requiring a double-strand break or a distinct template.

Promoter: DNA region necessary to initiate proximal transcription by binding RNA polymerases, general transcriptional machinery and transcription factors.

Regulatory sequences: Transcription factor-binding sites within regulatory regions that mediate gene regulatory function of regulatory regions or elements.

RNAi: RNA interference is a process in which transcripts are targeted by small complementary RNA, which eventually leads to transcript degradation and gene knockdown.

Silencer: A regulatory DNA region that binds repressors to decrease the likelihood of transcription.

TADs: Topologically associating domain, a region of the genome in which the contained sequences are more likely to physically interact with each other.

Topological elements: Boundary elements which constrain the physical interaction of chromosomal regions, such as TADs.

---

gene at will. This ability has already led to the discovery of new paradigms in gene regulation and, as new CRISPR-based tools and methods continue to be developed and perfected, promises to lead to even more insights into how genes are regulated and cells are specified during development. Here, we introduce the different tools and discuss how they can be applied to further our understanding of gene regulation in development and differentiation. We then discuss how CRISPR-based tools can be employed to precisely control gene expression in experimental systems and disease therapeutics.

The genome editing toolbox

The CRISPR genome editing toolbox has expanded substantially over the last few years since the first demonstration of CRISPR-Cas9-nuclease genome editing in mammalian cells (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013). Adapted from a bacterial immune system, the CRISPR system has been co-opted for use in mammalian genome editing by engineering single guide RNAs (sgRNAs) that complex with a CRISPR-associated nuclease protein, most commonly Cas9, to induce double-stranded breaks (DSBs) at a specific location in the genome, specified by the sequence of the sgRNA and a neighboring protospacer-adjacent motif (PAM; see Glossary, Box 1) (Fig. 1; Fig. 2A).

Improvements to CRISPR-Cas specificity
CRISPR-Cas targeting

Initially, CRISPR editing was restricted to sequences proximal to the PAM. Now, however, a plethora of Cas9 and Cas12a variants from distinct bacteria and phages have been characterized, each with

distinct PAMs (reviewed by Manghwar et al., 2019; Pausch et al., 2020). In addition, Cas9 variants have been engineered with 'looser' PAM recognition that serve to broaden the targeting range (Kleinstiver et al., 2015a,b; Nishimasu et al., 2018; Walton et al., 2020). Furthermore, variants with decreased off-target editing have been developed, although they concomitantly reduce editing efficiency (Kim et al., 2020; Kleinstiver et al., 2016; Slaymaker et al., 2016). Cas9-nucleases now exist to target nearly any PAM sequence with near-complete editing efficiency in most cell types.

Mutation prediction

The spectrum of insertion/deletion mutations (indels) generated by *Streptococcus pyogenes* SpCas9-nuclease, the most widely used and well-characterized Cas9 protein, is highly predictable given the target sequence and cell type. Therefore, an appreciable fraction of genomic targets yield precise outcome distributions dominated by a single indel product (Allen et al., 2019; Shen et al., 2018). There is now a convenient webtool (http://indelphi.giffordlab.mit.edu/) capable of predicting the outcome distribution for any sgRNA, and the ability to predict outcome distributions enables the design of sgRNAs that maximize frameshifts or disruption of genomic features of interest (Shen et al., 2018).

Prime editing and base editors

Prime editing (see Glossary, Box 1) employs a fusion of Cas9-nickase and a reverse transcriptase enzyme to enable precise genotypic replacement and short insertions (≤44 nt) or deletions (≤80 nt) with a substantially improved ratio of desired to undesired
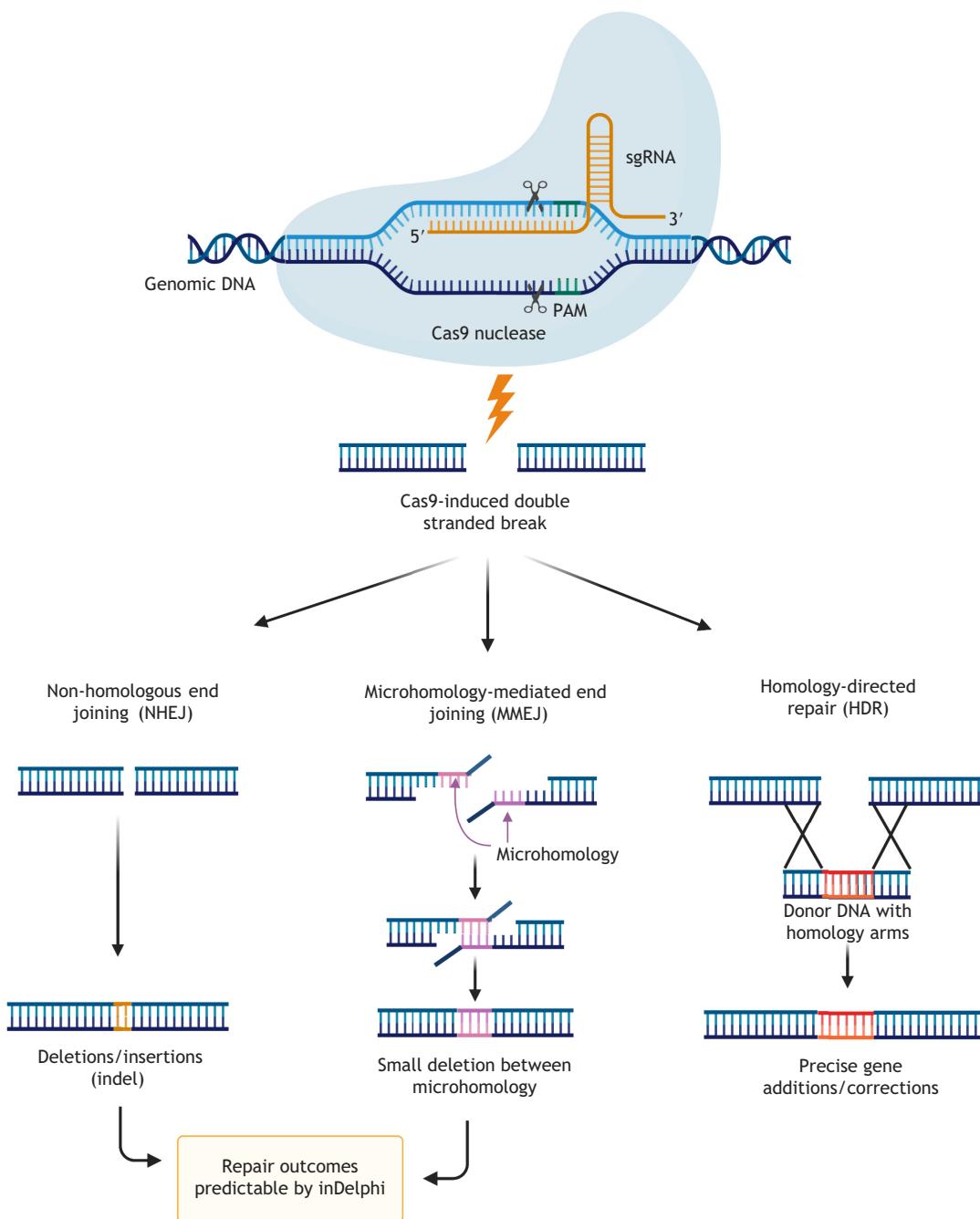
Fig. 1. Cas9 nuclease repair outcomes. Cas9-nuclease, which is guided to the target DNA region by a single guide RNA (sgRNA) because of complementarity between the sgRNA and the target DNA, creates double-stranded DNA breaks (DSBs). DSBs are imperfectly fixed by endogenous host DNA repair mechanisms such as non-homologous end joining (NHEJ) and microhomology-mediated end joining (MMEJ), creating short insertions/deletions (indels) around the cleavage site. If a DSB is created in the presence of an external DNA fragment bearing homology arms flanking the DSB site, this external DNA can integrate into the targeted position of the genome through cellular homology-directed repair (HDR). CRISPR-induced repair outcomes for any given sgRNA are predictable by the algorithm inDelphi (http://indelphi.giffordlab.mit.edu/; Shen et al., 2018), enabling selection of sgRNA designs that maximize desirable edited products.

editing genotypes (Anzalone et al., 2019) (Fig. 2B). However, prime editing currently has low editing efficiency (<30% of alleles) and often requires substantial optimization of each targeting strategy. Base editing (see Glossary, Box 1), in which Cas9-nickase is fused to an enzyme capable of chemically transforming nucleotide identity, has similarly been shown to mediate nucleotide transitions with high efficiency and specificity (Gaudelli et al., 2017; Komor et al., 2016) (Fig. 2B).

Alternative applications

In addition to CRISPR-Cas9 mutagenesis, CRISPR-Cas9 has also been used to epigenetically modify genomic regions through the fusion of nuclease-dead Cas9 (dCas9; see Glossary, Box 1) variants with epigenetic modifying agents (reviewed by Pickar-Oliver and Gersbach, 2019), including epigenetic silencers (CRISPRi; Fig. 2C) and transcriptional activators (CRISPRa; Fig. 2D) (see Glossary, Box 1).
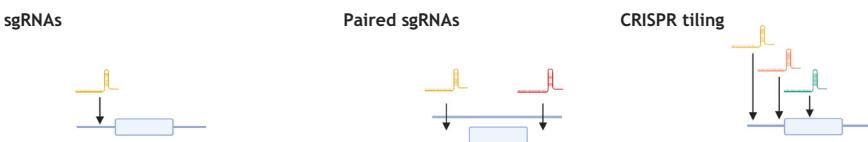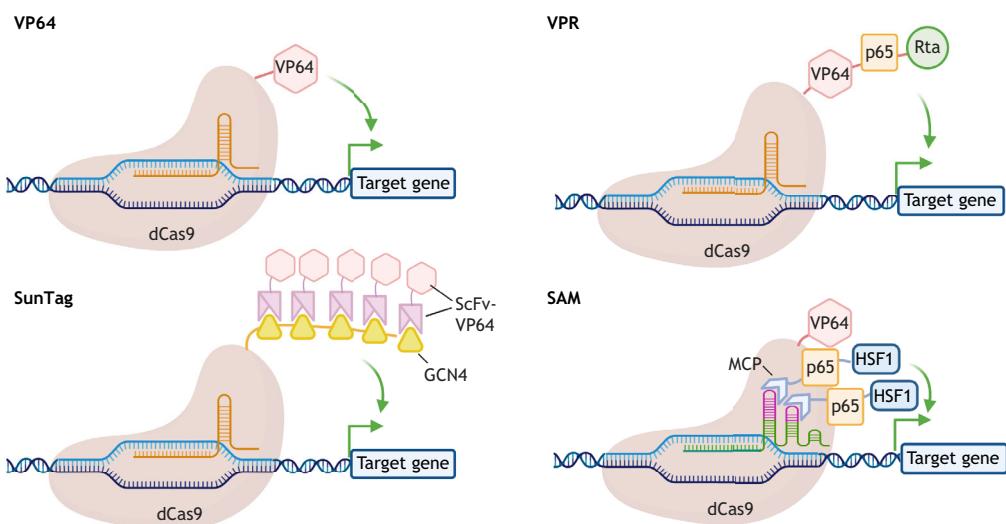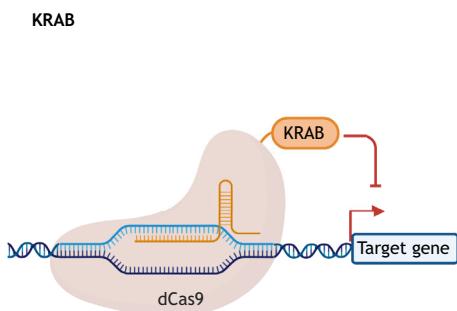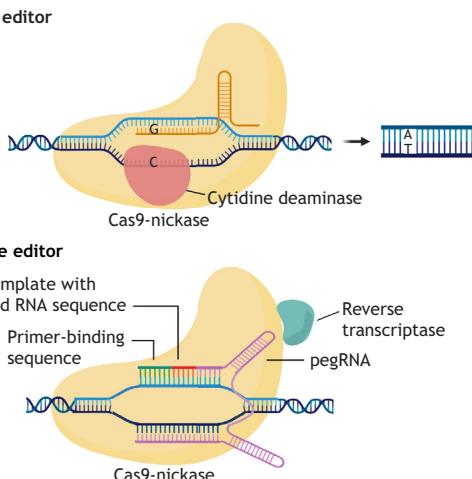
Fig. 2. The CRISPR toolbox. (A) Strategic CRISPR target designs enable a range of desired genome edits. One single guide RNA (sgRNA) targeting the coding region of a gene of interest can be used to induce a frameshift and thus a truncated nonfunctional protein product. Use of a pair of sgRNAs can excise an entire region of interest. Tiling sgRNA libraries can be used in pooled screens to map functional regions within coding and non-coding DNA. CRISPR-based tools with amino acid modifications and/or effector peptide/protein additions have been developed to perform distinct editing functions. (B,C) CRISPR activation (CRISPRa) (B) and CRISPR interference (CRISPRi) (C) enable epigenetic modification through use of a nuclease-deficient Cas9 (dCas9) enzyme in which two amino acids (D10A and H840A) in catalytic domains (RuvC and HNH) are mutated. In CRISPRi and CRISPRa, dCas9 is attached to transcriptional silencer or epigenetic activators domains, respectively, enabling modulation of the expression of target genes when recruited to the promoter or enhancer region. Various versions of CRISPRi and CRISPRa tools have been devised. First generation CRISPRa tools used direct attachment of a single activation domain such as the *Herpes simplex* tandem activator protein VP64. Among second-generation CRISPR activators, VPR entails attachment of three distinct activation domains [VP64, p65 and Epstein-Barr virus R transactivator (Rta)] to dCas9, SunTag entails indirect attachment of 10 VP64 peptides to dCas9 through binding of a single chain variable fragment (scFv) antibody to a tail of dCas9-fused peptide ligands, and SAM combines a direct dCas9-VP64 and recruitment of 2 activator domains [p65 and heat shock factor 1 (HSF1)] to sgRNA-embedded MS2 RNA aptamers (2× MS2). (D) Prime editors and base editors are capable of precise genome editing without DSBs. Prime editors catalyze short sequence replacements, insertions or deletions through use of a Cas9-nickase fused to a reverse transcriptase (RT) enzyme. The sgRNA used by prime editors is extended with a sequence that acts as an RT primer and template, stimulating DNA synthesis of the nicked DNA strand, which is integrated through cellular DNA repair mechanisms. Base editors are composed of a Cas9-nickase fused to a nucleotide deaminase enzyme, which catalyzes specific base transitions (C to T for cytidine deaminases and A to G with TadA adenine deaminase). MCP, MS2 coat protein; pegRNA, prime editing guide RNA.

CRISPRi

dCas9 can modestly inhibit transcription in human cells when directed to the proximal promoter or coding region of a gene by sterically hindering the transcriptional machinery or RNA polymerase processivity, respectively (Qi et al., 2013). To further enhance the efficiency of repression in mammalian cells, various transcriptional repression domains have been tethered to dCas9 among which KRAB (Krüppel-associated box) domains provide the most consistent and effective transcription repression (Gilbert et al., 2013). dCas9-KRAB efficiently silences transcription when targeted to the promoter (see Glossary, Box 1) (Gilbert et al., 2013) and/or enhancer (see Glossary, Box 1) (Thakore et al., 2015; Gao et al., 2014) of a target gene, and CRISPRi has been shown to provide consistent repression across most genes in the genome (Gilbert et al., 2014; Horlbeck et al., 2016).

CRISPRa

Ectopic gene activation has been achieved by recruiting transcriptional activator domains through direct or indirect fusion to dCas9. Initial CRISPRa platforms used direct dCas9 fusion of four copies of the *Herpes simplex* virus protein VP16 (VP64) (Gilbert et al., 2013; Perez-Pinera et al., 2013; Maeder et al., 2013; Farzadfard et al., 2013). As these first generation CRISPRa platforms gave weak and inconsistent gene activation, different activator domains have been fused together and recombined to dCas9 either directly or indirectly through antibody and RNA-binding protein intermediaries (known as second generation CRISPRa systems), significantly improving the activation of target genes (Chavez et al., 2015, 2016; Konermann et al., 2015; Tanenbaum et al., 2014) (Fig. 2D). Nonetheless, CRISPR-based gene activation is still inconsistent and weak at many loci.

CRISPR-based regulation of RNA

Finally, the last few years have seen rapid advances in the use of RNA-modulating CRISPR-based approaches (reviewed by Smargon et al., 2020). Such approaches enable cleavage, editing, fluorescent tracking and pulldown of specific transcripts. In particular, the advent of RNA-modulating CRISPR enzymes of the Cas13 family (O'Connell, 2019) promises to improve our understanding of the roles of RNA in gene regulation. Cas13 enzymes have been shown to cut RNA (Abudayyeh et al., 2017), edit adenosine to inosine in RNA (Cox et al., 2017) and image RNA (Yang et al., 2019). Altogether, since the first demonstration of CRISPR-Cas9 genome editing in mammalian cells (Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013), a powerful and varied toolbox has been constructed, facilitating precise and scalable manipulation of the genome.

Understanding gene regulation
Dissecting gene function

Understanding the genetic underpinnings of cell fate acquisition during development has been a longstanding goal in developmental biology. Traditionally, gene function has been unraveled based on phenotypic data from forward and reverse genetic screens in model organisms. For example, projects such as the International Mouse Phenotypic Consortium have now generated knockouts of over 6000 genes in mice (Dickinson et al., 2016; Cacheiro et al., 2019), largely using techniques that predate CRISPR-Cas9, such as ENU mutagenesis. Such organism-wide and tissue-specific knockouts of TFs have complemented techniques, such as chromatin immunoprecipitation sequencing (ChIP-seq), that measure TF binding to reveal roles for hundreds of TFs in organismal

development (Spitz and Furlong, 2012). CRISPR-Cas9 has largely replaced traditional methods in the production of transgenic and knockout animals given its rapid timeline and high editing efficiency (Burgio, 2018). More transformatively, CRISPR-Cas9 has expanded the scope of efforts to understand how gene networks control cell fate acquisition.

In addition to animal models, human pluripotent stem cells (hPSCs), which can differentiate into any human lineage, provide an *in vitro* model system to dissect the genetic underpinnings of human cell fate acquisition. Early studies employed RNAi-based loss-of-function (RNAi; see Glossary, Box 1) screens to dissect the transcriptional programs driving hPSC specification (Hu et al., 2009). The ease, precision and scalability of CRISPR-Cas9 has improved upon this screening approach (Fig. 3A). Such screens have used flow cytometric sorting in order to identify sgRNAs that enhance or inhibit specific differentiation outcomes, either through the introduction of a GFP reporter transgene or through the use of antibodies that report on expression of a lineage-specific gene (Li et al., 2019; Parnas et al., 2015; Xu et al., 2020). Through this approach, genome-wide CRISPR-Cas9 knockout screens (see Glossary, Box 1) have dissected genetic underpinnings of human pluripotency states (Fu et al., 2019; Li et al., 2018), as well as revealing novel roles for JNK/Jun signaling in endoderm differentiation and ZIC2 in cardiac differentiation (Li et al., 2019; Xu et al., 2020).

Directed differentiation paradigms also allow the interrogation of later stages of differentiation. To this end, a doxycycline (dox)-inducible Cas9 (iCas9) system has been developed to induce gene knockout at any chosen time point, which allows investigation into genes specifically required at distinct stages of differentiation (Boyle et al., 2017; Khera et al., 2018; Zhu et al., 2016) (Fig. 3A). In addition to studying particular differentiation paths, CRISPR-Cas9 can be combined with emerging gastruloid models (Moris et al., 2020) to study the genetic underpinnings of human early embryonic patterning in a scalable and representative model system, which has traditionally been difficult (Shahbazi, 2020). The flexibility of CRISPR-Cas9 approaches to introduce genetic perturbations at scale at defined developmental stages opens the door to genetic dissection of human developmental processes that were previously impervious to such probing.

Although marker gene readouts are limited in what they reveal about the causes of aberrant differentiation, a revolution in the ease, throughput and affordability of single-cell RNA-sequencing (scRNA-seq) opens the door to unraveling gene networks in finer detail. By linking single-cell transcriptomes to corresponding sgRNA perturbations, it is now possible to perform CRISPR-Cas9 screens using scRNA-seq as the readout (Adamson et al., 2016; Dixit et al., 2016; Jaitin et al., 2016) (Fig. 3A). Thus, instead of monitoring single marker genes, the response of entire gene regulatory networks (GRNs; see Glossary, Box 1) to genetic perturbations can be analyzed in high-throughput. Such approaches have been employed to group gene functions by the pathway in which they act, to refine regulatory networks by comparing transcriptomic phenotypes for a group of gene knockouts, and to understand defects in stem cell differentiation (Adamson et al., 2016; Dixit et al., 2016; Genga et al., 2019; Jaitin et al., 2016). In addition, dual-sgRNA screening, in which combinations of multiple sgRNA-induced perturbations are introduced in the same cell, have been applied to investigate how genes interact with each other (epistatic relationships; see Glossary, Box 1) to alter cell state (Adamson et al., 2016) (Fig. 3A). This multiplexing capability is crucial to dissect the complicated and often redundant functions of genes within signaling pathways.

DEVELOPMENT

**A Dissecting developmental gene and pathway function**
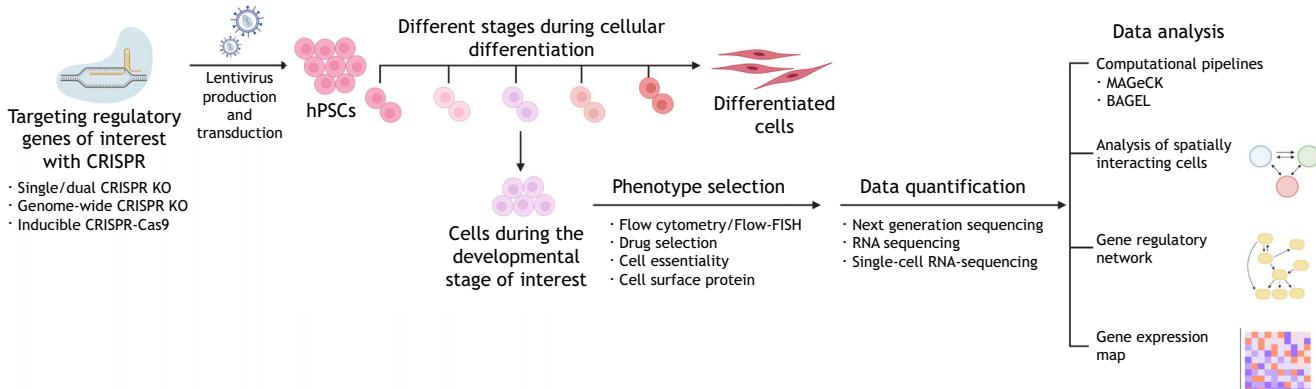
B Identifying and characterizing regulatory regions
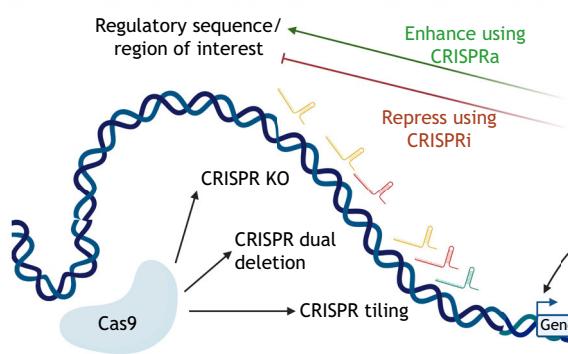
C Post-transcriptional regulation of gene expression and regulatory RNAs
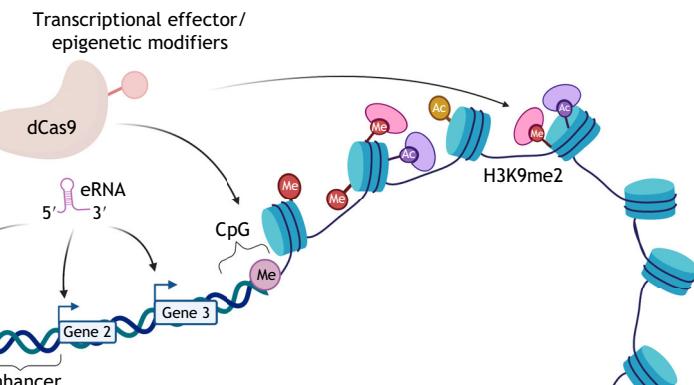
D Dissecting the roles of epigenetic modifications in gene expression
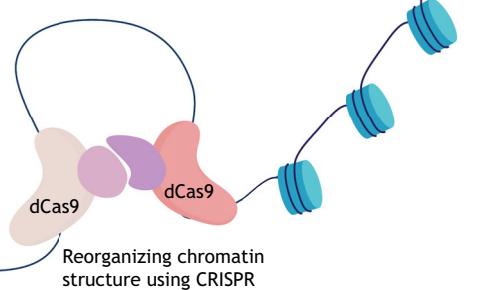
E Understanding 3D genomic topology
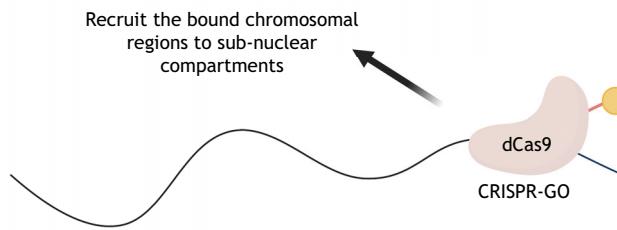
Fig. 3. Applications of the CRISPR toolbox to elucidate gene regulation. (A) CRISPR-based gene loss-of-function screens, through permanent or inducible mutation strategies, have eased the investigation of genetic underpinnings during different stages of cellular development. A generic screen paradigm involves the use of a pooled single guide RNA (sgRNA) library, transduced into cells using lentivirus, to target all or a focused set of genes/genomic elements of interest. Relative representation of each sgRNA can be quantified before and after phenotypic selection using techniques such as cell survival/proliferation under defined conditions or flow cytometry-based cell sorting, followed by sgRNA quantification methods including next-generation sequencing or single-cell RNA-sequencing, to investigate genes/elements involved in developmental processes. (B-D) Cas9 enzyme variants have facilitated elucidating gene regulation. Cas9 can be employed for single gene/element mutation, deletion and tiling of gene regulatory sequences. dCas9 fused with a transcriptional effector domain has been used to enhance or repress regulatory regions to understand their involvement in gene regulation and target epigenetic factors and modifications such as histone modifications, CpG methylation and enhancer transcripts to dissect their role in regulating gene expression. (E) The effect of 3D chromosomal architecture on gene expression can also be investigated with CRISPR tools that bring targeted chromosomal regions into proximity or allow recruitment to distinct sub-nuclear regions.

However, the quality of data provided from current scRNA-seq approaches is heterogeneous; unique transcripts can be exceedingly low for lowly-expressed genes (such as many cell fate-determining TFs) (Zhang et al., 2019a), and approaches such as CRISPR-Cas9-nuclease targeting do not produce uniformly disruptive mutations in every cell (Shi et al., 2015). As a result, proper computational analysis that accounts for such noise and dropout is crucial (Luecken and Theis, 2019), and scRNA-seq is currently best-suited to applications for which phenotypes involve large changes to the transcriptome, not changes in expression of individual genes. We anticipate that the pairing of CRISPR-Cas9 and scRNA-seq will shed light on key questions such as why defined stem cell differentiation paradigms sometimes yield non-homogeneous end products (Veres et al., 2019), how genetic mutations in the same set of developmental genes cause a range of phenotypic severity from organ agenesis to incompletely penetrant adult-

onset disease (Ashcroft and Rorsman, 2012), and what molecular mechanisms underlie common disease-associated genetic variants (Box 2).

Understanding GRNs

Developmental patterning is driven by cell-cell communication driving *cis*-regulatory binding changes that alter expression of TFs that further alter cellular gene expression and function. To study cell-cell communication, a set of emerging single-cell technologies has unlocked other phenotypic measurements that can be productively paired with CRISPR-Cas9 screening (Fig. 3A). In PIC-seq (see Glossary, Box 1), physically interacting cells (PICs) are sorted by flow cytometry and their combined RNA transcriptomes are sequenced and compared with their single-cell transcriptomes. Computational modeling is then used to map interactions between cells (Giladi et al., 2020). This method could be combined with perturbing signaling pathways to determine how intercellular signaling impacts cell-cell interactions, for example in Notch pathway-driven processes (Henrique and Schweisguth, 2019) and in stem cell-niche interactions (Centonze et al., 2020; Chacón-Martínez et al., 2018).

To investigate changes in *cis*-regulatory binding, droplet single-cell assay for transposase-accessible chromatin using sequencing (dscATAC-seq) can be used to measure accessible chromatin at single-cell resolution (Lareau et al., 2019). If combined with CRISPR, this method could be used to investigate the role of altering signaling pathways, TFs and *cis*-regulatory elements on chromatin accessibility in high-throughput. The combination of CRISPR-Cas9 screening and sensitive single-cell measurement promises to elucidate such networks in greater detail than has been possible (Davidson, 2010).

Box 2. The role of CRISPR-Cas9 in elucidating disease-associated genetic variation

Most common diseases (e.g. cancer, heart disease, autoimmune disease, type II diabetes) have a substantial genetic risk component. Recent research has converged on the finding that such diseases are more often caused by a combination of multiple low impact variants (polygenic risk) rather than one high-effect variant alone, as is the paradigm in rare monogenic diseases (Boyle et al., 2017; Khera et al., 2018; Nishizaki and Boyle, 2017). In order to identify genetic variants associated with risk of complex diseases, one common tool is the genome wide association study (GWAS) (Tam et al., 2019), in which each variant is tested for correlation with disease in a large cohort of often >100,000 people. Although such screens identify dozens to hundreds of genomic loci associated with common diseases (e.g. Crohn's disease, coronary artery disease, type II diabetes), it can be difficult to interpret disease-associated variants (daVs) found from GWAS given that >90% of daVs are in non-coding regions, so it is difficult to connect daVs to impacted genes. In addition, co-inherited daVs can be in strong linkage disequilibrium, and therefore there are often multiple linked variants that correlate equivalently to disease, complicating identification of causal daVs through GWAS data alone. As a group, daVs are enriched in *cis*-regulatory elements, indicating that these variants often act through regulating gene expression (Gallagher and Chen-Plotkin, 2018). CRISPR-Cas9 is a useful tool in deciphering the function of GWAS variants. CRISPR-Cas9 tools can be used to install candidate daVs to study their effects on gene regulation through functional assays such as ATAC-seq, CHIP-seq, RNA-seq, protein expression, or molecular and cellular assays (Smith et al., 2018). Understanding the gene regulatory functions of daVs is a frontier in understanding the etiology of and proposing therapies for common genetic diseases.

Identifying and characterizing regulatory regions
Tiled CRISPR screens

Our ability to predict how changes to non-coding regions will impact cellular gene expression and function is limited. As non-coding variation within and between species is a major driver of disease and evolution, respectively (Banerjee and Sherwood, 2017; Nishizaki and Boyle, 2017; Reilly and Noonan, 2016; Tickle and Urrutia, 2017), it is imperative to improve our interpretation of the regulatory genome.

To address these questions, a new class of high-throughput assay has been developed in which CRISPR sgRNAs non-specifically tile or target specific candidate regulatory regions, followed by measurement of resulting changes in gene expression (Canver et al., 2015, 2017; Diao et al., 2016, 2017; Fulco et al., 2016, 2019; Gasperini et al., 2019; Klann et al., 2017; Korkmaz et al., 2019; Rajagopal et al., 2016) (Figs 2A and 3B). These approaches provide a platform to assess whether putative regulatory regions truly cause changes in gene expression. Such assays vary in the type of CRISPR-based manipulation induced at the target site and the readout method, tailoring the assays to measure distinct activities.

For example, Cas9-nuclease tiling assays enable pinpointing of regulatory regions within ∼20 nucleotide regions because this is the size of a typical indel (Canver et al., 2015; Korkmaz et al., 2016; Rajagopal et al., 2016; Sanjana et al., 2016). This fine resolution is helpful to resolve effects of individual binding motifs; however, indels induced by a given sgRNA are variable (Shen et al., 2018; van Overbeek et al., 2016), leading to a high degree of variability in the phenotypes that can result from the activity of each sgRNA. Paired sgRNAs can be used to create defined deletions, excising candidate regulatory regions (Diao et al., 2017), which has a benefit of precisely defining the bounds of the candidate region with the downside that Cas9-induced deletions often do not constitute the majority of edited products (Figs 2A and 3B). In addition, CRISPRi provides a tool to perform uniform epigenetic repression because, in contrast to Cas9-nuclease, a given sgRNA should induce the same phenotype in every cell (Fulco et al., 2016; Gasperini et al., 2019) (Fig. 3B). However, CRISPRi (see Glossary, Box 1) has a resolution limit of ∼200-500 bp because of the spreading of epigenetic repression, preventing dissection of the role of sequence features within regulatory elements (Fulco et al., 2019).

Conversely, CRISPRa screening methods allow probing of genomic regions capable of acting as distal enhancers (Klann et al., 2017) (Fig. 3B), but the generalizability of CRISPRa effectors is not well-established, warranting caution in the interpretation of such data. Likewise, several methods for evaluating gene expression phenotypes have been used, including cell survival and drug resistance when targeting regulatory regions associated with genes essential to such traits, fluorescent reporter gene expression, cell surface protein expression, RNA fluorescence *in situ* hybridization (FISH)-based FlowFISH and scRNA-seq (Fig. 3A).

Data analysis

Several computational pipelines have been developed to analyze data from tiled CRISPR screens. Tiled non-coding CRISPR screens differ from gene-targeting CRISPR screens based on the assumptions that can be made about expected outcomes. In gene-targeting screens, it is expected that sgRNAs that target the same gene will yield identical phenotypes. This assumption is not entirely true in practice because sgRNAs can differ in their editing efficiency (Kim et al., 2019b), their induction of frameshifts (Shen et al., 2018) and the importance of the targeted region of the gene (Shi et al., 2015); nonetheless, computational pipelines have been purpose-

built to analyze gene-targeting CRISPR screening data effectively, including MAGeCK (Li et al., 2014) and BAGEL (Hart and Moffat, 2016). A distinct set of assumptions is required to facilitate interpretation of tiled non-coding CRISPR screen data. Unlike in gene-targeting screens, there are no sets of sgRNAs that can be assumed to yield identical phenotypes; instead, partial phenotypic correlation is expected among neighboring sgRNAs that target the same regulatory element. However, where such elements begin and end in the genome and how much correlation to expect among different sgRNAs targeting a given element are not well known. Moreover, the expected phenotypic correlation depends strongly on the editing mode. CRISPRi tends to induce relatively uniform epigenetic repression of an entire regulatory element, whereas Cas9-nuclease can yield variable phenotypes at neighboring sgRNAs because neighboring sgRNAs may fail to disrupt the same TF-binding motif (Hsu et al., 2018; Rajagopal et al., 2016). With these parameters in mind, two pipelines, CRISPR-SURF (Hsu et al., 2018) and RELICS (Fiaux et al., 2020), have been designed to analyze tiled non-coding CRISPR screens. These platforms account for the expected partial correlation among neighboring sgRNAs and allow semi-supervised adjustment of the size of genomic regions that are expected to share phenotypic outcomes.

Insights from tiled CRISPR screens

These methods have allowed for the identification of new types of regulatory regions and the design of enhanced predictive models of regulatory function. For example, several such screens have converged on the regulatory importance of the promoters of neighboring genes (Fulco et al., 2016; Rajagopal et al., 2016), complementing computational studies that demonstrate widespread co-expression of neighboring genes (Mihelčić et al., 2019). In addition, a number of studies have shown that disrupting CTCF TF-binding sites through CRISPR-Cas9 editing is able to reconfigure long-range chromatin looping of entire chromosomal territories (Guo et al., 2015; Khoury et al., 2020; Korkmaz et al., 2019; Tarjan et al., 2019). This evidence converges on the idea that gene expression may be more influenced by shared regulatory regions within chromosomal neighborhoods than previously believed.

In addition, such screens have refined prediction of which regulatory elements are likely to influence the expression of a nearby gene and to what extent. By analyzing thousands of enhancer-gene contacts, Fulco and colleagues developed the activity by contact (ABC) model, which calculates the quantitative effect of a regulatory element on expression of a nearby gene based on the strength of the region as an enhancer (measured by chromatin accessibility and H3K27 acetylation) and estimates contact frequency between the regulatory element and target gene promoter in that cell type (determined by Hi-C, a technique that measures three-dimensional architecture of chromatin) (Fulco et al., 2019). Although predictions are highly cell-type specific and rely on additional epigenomic and Hi-C data that is not available for most cell types, the ABC model is useful for predicting the relevance of genetic variants in non-coding regions. A high-throughput enhancer-targeting CRISPRi screen that used scRNA-seq as readout similarly found that proximity to the target gene and strong Hi-C contact frequency with the target promoter are correlated with enhancer activity (Gasperini et al., 2019).

Future directions

Perturbational approaches promise to further advance our understanding of gene regulation. Ultimately, a predictive model of gene expression should enable prediction of the level of

expression of a gene, based solely on the adjacent regulatory genomic sequence and the gene expression profile of a cell. In the ABC model, the difficult parts of such modeling are skipped over by taking into account epigenomic data and contact frequency, which integrate binding of typically dozens of TFs at a regulatory region (Partridge et al., 2020; Sherwood et al., 2014). How such factor combinations lead to transcriptional activation (Smith et al., 2013), how distal regulatory elements interact in sometimes specific ways with neighboring promoters presumably because of specific TF-mediated interactions (Zabidi et al., 2015) and how multiple distal regulatory elements interact to contribute to expression of a given gene remain outstanding challenges in the field.

A recent CRISPRi enhancer tiling study has made the intriguing observation that epigenetic inhibition of pairs of enhancers adjacent to the same gene provides more repression than would be expected from adding the repressive effects of each individual enhancer (Xie et al., 2017). This analysis was only performed at two genes, so it remains to be seen whether non-linear combinatorial activation is the exception or the rule for enhancer function. We anticipate that CRISPR-based tools will contribute greatly to understanding these questions, eventually leading to a more specific lexicon of gene regulation. Instead of merely grouping regulatory elements into rough classes such as enhancers, promoters and silencers, we anticipate that a much more rich understanding of how specific TFs and combinations thereof orchestrate lineage- and stimulus-specific transcriptional dynamics will emerge.

To improve predictive resolution from the level of kilobase-scale regulatory elements to the level of individual <10 nucleotide TF-binding sites requires editing approaches with nucleotide-level precision. Pairing tiled Cas9-nuclease screening with a computational model such as inDelphi (Shen et al., 2018) that predicts the range of indels created by Cas9-nuclease editing would improve our understanding of which genomic elements are likely to be altered by each sgRNA. Exploration of screens using other Cas species such as Cas12a and Cas3 that produce differently sized indels could also be worthwhile (Dolan et al., 2019; Morisaka et al., 2019; Zetsche et al., 2015). In addition, precise CRISPR targeting methods, such as base editing and prime editing (Anzalone et al., 2020), promise to improve our ability to monitor the regulatory consequences of precise genome modifications.

In addition to CRISPR-based approaches that alter the native genomic sequence, massively parallel reporter assays (MPRAs) provide an important complementary approach to dissect regulatory element function. MPRAs are widely used to test the transcriptional regulatory activity of thousands of candidate regulatory sequences in parallel through linking such regulatory activity with the transcription of unique barcode sequences (Arnold et al., 2013; Inoue and Ahituv, 2015; Melnikov et al., 2012). MPRAs have been used to identify differences in transcriptional output associated with pathogenic genomic variants (Tewhey et al., 2016; van Arensbergen et al., 2019). Two caveats of MPRAs is that they have traditionally been performed on plasmids, which lack the native chromatin environment of the genome, and they have typically assessed transcriptional output but not the underlying mechanistic components of gene regulation. CRISPR-based homology-directed repair (HDR) provides solutions to both of these issues, enabling study of large numbers of variant regulatory elements in defined genome-integrated settings. Using CRISPR-based HDR, thousands of synthesized ~100 bp DNA sequences can be integrated at a fixed locus in mouse embryonic stem cells and the binding strength of a TF at each sequence can then be measured. Recent work

has demonstrated the feasibility and power of this approach: one study has found that the binding of the TF Tcf7l2 depends on the local level of chromatin accessibility provided by cofactors, as well as whether Tcf7l2 binds on the same side of the DNA helix as its cofactors Oct4 and Klf4 (Szczesnik et al., 2020). Another recent study used a similar technique to measure changes in chromatin accessibility induced by the integration of thousands of 100-bp genomic sequences, revealing that such short sequences are sufficient to induce differential accessibility between stem cells and differentiated endoderm (Hammelman et al., 2020). Sequence determinants of TF binding and chromatin accessibility can now even be measured at single-allele resolution through marking accessible bases with adenine methylation and performing PacBio long-read sequencing (Stergachis et al., 2020). This, in combination with measurements of transcriptional activation, will provide a powerful complement to CRISPR screens to dissect the sequence features governing regulatory element activity and how such elements are influenced by disease-associated non-coding variants found in genome-wide association studies (GWAS; see Glossary, Box 1; Box 2).

Post-transcriptional regulation of gene expression and regulatory RNAs

CRISPR-Cas9 tools have also been used to explore roles of non-coding RNAs and have shed light on the role of enhancer transcripts (eRNAs; see Glossary, Box 1), which are known to be produced from most active enhancers (Kim et al., 2010) (Fig. 3C). In spite of their abundance, it has been controversial whether eRNAs are causally important to gene transcription or whether they are inert byproducts of the recruitment of RNA polymerase to gene regulatory loci (Andersson et al., 2014; Henriques et al., 2018; Young et al., 2017). A recent study has found that loss of the RNA Polymerase II elongation factor *SPT5* (also known as *Supt5* or *Supt5h*) coordinately decreases the expression of a subset of genes and eRNA transcripts at linked enhancers. This eRNA loss is causal because when the CRISPRa tool dCas9-VPR is recruited to a distal enhancer in *SPT5*-depleted cells, it drives eRNA transcription that restores neighboring gene expression (Fitz et al., 2020). This finding has been supported by another recent study showing that eRNA induction through distal enhancer-targeted CRISPRa upregulates local gene expression in a breast cancer model (Zhang et al., 2019b). The authors suggest that eRNA transcription facilitates enhancer-promoter interaction at pre-activated enhancers, although it is currently difficult to disentangle eRNA-mediated effects from distal gene activating effects of CRISPRa recruitment. Recent work has bolstered the idea that eRNAs enhance gene activation by showing that low-level RNA production, as is produced by enhancer transcription, increases the formation of transcription-promoting condensates, which are subsequently dissolved by robust gene transcription (Henninger et al., 2021). Although enticing, none of the evidence presented above is conclusive, and CRISPR-based approaches promise to play an outsized role in deciphering the roles, if any, of eRNAs.

In addition to elucidating roles of eRNAs, RNA-targeting CRISPR approaches promise to: (1) clarify functions of long intergenic non-coding RNAs (lincRNAs) (Engreitz et al., 2016; Kopp and Mendell, 2018); (2) study roles of RNA motifs and ribonucleoprotein complexes at specific DNA targets (Shechner et al., 2015); (3) dissect steps involved in the transcription process (Wang et al., 2019b).

Dissecting the roles of epigenetic modifications in gene expression

Epigenetic factors, such as post-transcriptional modifications of histones, DNA methylation and non-coding RNAs, contribute to cellular gene expression, differentiation potential and response to stimuli (Allis and Jenuwein, 2016). Dissecting the role of these epigenetic modifications on gene regulation has been accelerated by the ability to site-specifically alter the epigenome using the CRISPR/(d)Cas9 toolbox (Fig. 3D). dCas9 has been employed to recruit epigenetic modifiers to evaluate their functional consequences by being fused to epigenetic modifying enzymes and complexes including lysine-specific histone demethylase 1A (LSD1; KDM1A) (Kearns et al., 2015), histone acetyltransferase p300 (Hilton et al., 2015), disruptor of telomeric silencing 1-like histone H3K79 methyltransferase (Dot1l), PR/SET domain 9 (Prdm9) (Cano-Rodriguez et al., 2016), enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2) (O'Geen et al., 2019), BAF nuclear assembly factor 1 (BAF), heterochromatin protein 1 (HP1; CBX5) (Braun et al., 2017), DNA methyltransferases (DNMTs), and tet methylcytosine dioxygenase 1 (TET1) (Liu et al., 2018). These studies have collectively begun to address which epigenetic modifications lead to stable epigenetic memory and which modifications are transient cues that are rapidly overridden. This is a young field; the power of controlled site-specific genomic recruitment of domains from TFs and epigenetic regulators promises to enable much finer dissection of their specific roles in epigenetic and transcriptional modulation.

Understanding 3D genomic topology

In addition to the DNA and RNA elements involved in gene regulation, the three dimensional (3D) structure of chromosomes is known to play a major role in which genes are expressed. However, our knowledge of how this architecture is controlled and how the 3D genome impacts gene regulation is limited. CRISPR tools now enable manipulation of genome architecture (Fig. 3E). For example, by building and using a CRISPR tool named CLOuD9 (inducible and reversible chromatin loop reorganization using nuclease-deficient Cas9 and components of plant hormone ABA signaling pathway) in which chromatin loops can be dynamically manipulated in living cells, recent work has demonstrated that chromatin looping in the proper biological context is sufficient to alter gene expression (Morgan et al., 2017). Similarly, a light-activated-dynamic-looping (LADL) system, in which two genomic regions anchored by gRNAs are brought into proximity via a light-induced dCas9 fusion protein, shows that ectopically inducing proximity between an enhancer and a promoter that normally do not interact leads to a modest increase in target gene expression (Kim et al., 2019a). Finally, an approach named 'CRISPR-GO' (CRISPR genome organization), in which chromosomal regions can be reversibly recruited to distinct sub-nuclear compartments, has found that recruiting gene regions to the nuclear periphery or to Cajal bodies is sufficient to reduce reporter gene expression (Wang et al., 2018). These approaches add an extra dimension to gene regulatory perturbation, and as their precision is increased, they promise to improve our understanding of how gene regulation is influenced by nuclear organization.

Manipulating gene regulation

CRISPR-Cas9 is not just a powerful tool to decipher gene regulation, it also provides a method to alter gene expression and cellular state in ways that can be therapeutically relevant. The first therapeutic implementation of CRISPR-Cas9 to alter gene regulation has now been shown to be effective: following the finding that an erythroid-specific distal enhancer regulating the *BCL11A* gene mediates fetal hemoglobin persistence (Bauer et al., 2013), a clinical trial has now shown that CRISPR-Cas9-mediated disruption of this enhancer in hematopoietic stem/progenitor cells, followed by transplantation of edited cells into patients with

β thalassemia or sickle cell disease, alleviates disease-associated clinical phenotypes (Frangoul et al., 2021). This landmark trial provides a proof-of-principle that CRISPR-based manipulation of gene regulation is more than just a research tool. CRISPR-based tools are being employed in several other settings to manipulate gene regulation in order to model and potentially treat human disease.

Modeling disease using pluripotent stem cells

Directed differentiation of induced pluripotent stem cells (iPSCs) to disease-relevant cell types has promised a revolution in understanding mechanisms of human genetic disease (Cherry and Daley, 2012). Although there have certainly been some groundbreaking insights gained through this approach (Rowe and Daley, 2019), one lesson that has emerged is that iPSCs derived from different patients – and even from the same patient – can have substantial heterogeneity at the genetic and transcriptional levels, which can make measurements between cohorts difficult (Kilpinen et al., 2017; Popp et al., 2018; Volpato et al., 2018). A complementary approach has emerged to use CRISPR-Cas9 to induce specific genetic alterations in otherwise identical iPSC genetic backgrounds (Musunuru, 2013; Musunuru et al., 2018). By specifically inducing focal genetic changes in pluripotent stem cell lines with an otherwise fixed genetic and epigenetic background, much of the noise associated with distinct genetic backgrounds and iPSC lines can be avoided. The development of base editing and prime editing techniques makes this an attractive approach to assess the causality of specific variants (or combinations thereof) with uncertain causality for particular disease phenotypes. It is important to recognize the strong disease-modifying effects of genetic background (Fahed et al., 2020) that merit inducing a given variant in several distinct iPSC backgrounds. Ideally, a bank of iPSCs with diverse genetic background and defined polygenic disease risk would be characterized and available for such experiments. It is also worth noting that inducing CRISPR-Cas9 alterations in a fixed background is much easier and more scalable than generating new iPSC cohorts for each new study, especially given the strong propensity for oncogenic and copy number alterations to occur during the iPSC derivation process (Kilpinen et al., 2017; Popp et al., 2018; Volpato et al., 2018). Overall, for each project aimed at modeling disease through iPSCs, it is worth weighing the benefits and drawbacks of installing variants through CRISPR-Cas9 as opposed to the more laborious process of generating new patient-specific iPSC cohorts.

Cellular reprogramming with CRISPRa

Cellular reprogramming, the conversion of one cell type into another, has primarily been achieved by overexpressing transgenes encoding TFs (Takahashi and Yamanaka, 2006), sometimes in combination with signaling molecules (Pagliuca et al., 2014) or knockdown of genes (Qian et al., 2020). Although this reprogramming paradigm has been successfully employed to produce distinct types of neurons and other lineages (Cherry and Daley, 2012; Vierbuchen et al., 2010), the pace at which new cell types have been reprogrammed has been slow due to the technical challenges associated with screening for and validating reprogramming cocktails. In theory, CRISPRa is ideally suited to this challenge owing to the ease of targeting any gene and multiplexing candidates in a scalable manner.

CRISPRa has been employed in iPSC and neural reprogramming paradigms (Black et al., 2016; Liu et al., 2018; Weltner et al., 2018). In the CRISPRa iPSC reprogramming work (Weltner et al., 2018), reprogramming efficiency has been increased through activating not only the known iPSC-promoting TFs (*POU5F1*, *MYC*, *KLF4*, *SOX2* and *LIN28A*), but also a repetitive EGA-enriched Alu-motif (EEA-motif), which is enriched in promoters of genes expressed during human embryo genome activation (Töhönen et al., 2015), thus revealing an added benefit of CRISPRa that cannot easily be replicated using transgenes. Nonetheless, in all such paradigms, CRISPRa-based reprogramming has remained substantially less potent than transgene-based reprogramming, probably because the levels of TF induction required for efficient reprogramming cannot be reached with current CRISPRa technology. Thus, efficient CRISPR-based reprogramming remains a tantalizing prospect that appears just beyond the horizon. In the near future, more robust CRISPRa technology, paired with emerging approaches to predict TF cohorts capable of reprogramming cellular function (Cahan et al., 2014; Rackham et al., 2016), promises to usher in a new age of accelerated derivation of reprogramming cocktails to match the discovery and categorization of cell types in the human body brought on by the scRNA-seq revolution (Macosko et al., 2015; Regev et al., 2017).

CRISPRa for therapeutic correction of haploinsufficiency

Finally, there is promise that, in addition to the immense promise of CRISPR-based gene editing in treating genetic disease (Dunbar et al., 2018), CRISPR-based epigenetic modulation could have future therapeutic applications. Although modular oligonucleotide-based gene therapy approaches to reduce gene expression, such as antisense oligonucleotides and RNAi, have already progressed to the clinic (Bennett, 2019), increasing gene expression has been more challenging. Gain-of-function gene therapy uses gene delivery, which currently cannot be used for genes that exceed the ~5 kb packaging limit of current state-of-the-art adeno-associated virus (AAV) vectors (Wang et al., 2019a). A recent study has shown that *in vivo* delivery of CRISPRa components targeted to gene promoters or enhancers is able to restore target gene levels enough to rescue the obesity phenotypes in mouse haploinsufficient genetic disease models induced by heterozygous mutations in the *Sim1* and *Mc4r* genes (Matharu et al., 2019). Importantly, this study has successfully employed *Streptococcus aureus* Cas9, a smaller Cas9 species that fits into a single AAV vector along with an sgRNA, enabling upregulation of genes through AAV vectors, the size of which exceeds the AAV packaging limit. It is worth mentioning the caveat that a substantial fraction of people have pre-existing immunity to Cas9 (Charlesworth et al., 2019) and Cas9 has been shown to be immunogenic (Chew et al., 2016; Wang et al., 2015), so long-term therapeutic expression of Cas9 may be inadvisable clinically.

Concluding remarks

It is remarkable how fast the CRISPR-Cas9 field has progressed since it was first demonstrated in genome editing 8 years ago. A panoply of CRISPR-based tools has emerged, capable not only of cleaving DNA and RNA, but of enabling precise site-specific base alteration, epigenetic modulation, imaging and sub-nuclear rearrangement.

It is no surprise that, as each new CRISPR-based tool has emerged, it has been readily co-opted to illuminate the enigmatic code used to generate an entire organism from a fixed genome (Table 1). Genetic screens in differentiating hPSCs are filling in the blanks of which genes control the formation of each organ, enriched by the emergence of scRNA-seq to classify mutant phenotypes and approaches to correlate such phenotypes with chromatin state and the identity of neighboring cells. The complex regulatory code that

Table 1. Summary of CRISPR-Cas tools

CRISPR-Cas tools	Function	Application(s)	Notes	Reference
CRISPR-Cas9 nuclease	DSB induction.	Generating KO cell lines/animals.	High efficiency, predictable outcomes.	Cong et al., 2013; Mali et al., 2013; Jinek et al., 2013
Prime editing	Precise genotypic replacement and short indel.	Precise genome editing; assessing the causality of specific variants.	Low editing efficiency.	Anzalone et al., 2019
Base editing	Base transitions.	Installing precise point mutations; assessing the causality of specific variants.	Higher efficiency than prime editing; only installs transition variants.	Komor et al., 2016; Gaudelli et al., 2017
CRISPR interference (CRISPRi)	Gene silencing.	Modulating gene expression; identifying regulatory regions.	Resolution limit of ~200-500 bp.	Qi et al., 2013; Gilbert et al., 2013
CRISPR activation (CRISPRa)	Gene activation.	Modulating gene expression; identifying regulatory regions; cellular reprogramming; therapeutic correction of haploinsufficiency.	Variable efficiency depending on gene target; distinct platforms required when targeting enhancers versus promoters.	Gilbert et al., 2013; Perez-Pinera et al., 2013; Tanenbaum et al., 2014; Konermann et al., 2015; Chavez et al., 2015
CRISPR-Cas9 KO screens	High-throughput sgRNA perturbations.	Discovering genes/regions with a specific function such as cell fate acquisition.	Functional readouts include flow cytometric sorting and cell survival. Can be temporally controlled via dox-iCas9.	Shalem et al., 2014; Wang et al., 2014
Dual-sgRNA screens	To induce deletions or combinations of perturbations.	Investigating epistatic interactions to alter cell state.	Deletions are lower efficiency than single sgRNA indels.	Adamson et al., 2016; Diao et al., 2017
CRISPR-Cas9+ scRNA-seq (e.g. Perturb-seq)	To link single-cell transcriptomes to corresponding sgRNA perturbations.	Grouping gene functions by the pathway in which they act, refining regulatory networks by comparing transcriptomic phenotypes for a group of gene KOs, understanding defects in stem cell differentiation.	Best-suited to applications where phenotypes involve large changes to the transcriptome, not changes in expression of individual genes.	Dixit et al., 2016; Adamson et al., 2016; Jaitin et al., 2016
CRISPR-Cas9 screens+PIC-seq	To map interactions between cells.	Determining how intercellular signaling impacts cell-cell interactions.	None.	Giladi et al., 2020
CRISPR-Cas9 screens+dscATAC-seq	To measure accessible chromatin at single-cell resolution after sgRNA perturbation.	Investigating the role of altering signaling pathways, transcription factors and cis-regulatory elements on chromatin accessibility.	None.	Lareau et al., 2019
Tiled non-coding CRISPR screens	To tile or target specific candidate regulatory regions and measure resulting changes in gene expression.	Identifying new types of regulatory regions.	Resolution and results vary depending on CRISPR tool.	Canver et al., 2015; Diao et al., 2016; Fulco et al., 2016; Rajagopal et al., 2016; Klann et al., 2017; Gasperini et al., 2019; Korkmaz et al., 2019
CRISPR-based HDR+MPRAs	To measure the binding strength of a TF at a large number of variant regulatory elements in a fixed locus.	Dissecting the sequence features governing regulatory element activity and how such elements are influenced by disease-associated non-coding variants.	None.	Szczesnik et al., 2020; Hammelman et al., 2020
Cas13 enzymes	To cut, edit or image RNA.	Investigating the roles of eRNAs, lincRNAs, RNA motifs, and ribonucleoprotein complexes; Dissecting steps involved in the transcription process.	None.	Abudayyeh et al., 2017; Cox et al., 2017; Yang et al., 2019
CLOuD9/LADL/ CRISPR-GO	To manipulate chromatin loops; To induce proximity between an enhancer and a promoter; To recruit chromosomal regions to sub-nuclear compartments.	Understanding how gene regulation is influenced by nuclear organization.	None.	Morgan et al., 2017; Kim et al., 2019a; Wang et al., 2018

CRISPR-GO, CRISPR-genome organization; DSB, double-stranded DNA breaks; dscATAC-seq, droplet single-cell assay for transposase-accessible chromatin using sequencing; HDR, homology-directed repair; indel, insertion/deletion; KO, knockout; LADL, light-activated dynamic looping; lincRNAs, long intergenic non-coding RNAs; MPRAs, massively parallel reporter assays; PIC-seq, physically-interacting cell sequencing; scRNA-seq; single-cell RNA-sequencing; sgRNA, single guide RNAs.

dictates gene expression dynamics is beginning to be illuminated, as gene regulatory elements and their underlying TF-binding motifs can now be altered in high-throughput, allowing inference of their normal function. In addition to alterations of regulatory DNA itself, local chromatin, non-coding RNA species and 3D genome organization can all be manipulated in ever more complex ways

through CRISPR-based tools. As a result, a more holistic view of the coordinated choreography involved in cell state transition is beginning to emerge.

As understanding of gene regulation improves, so does our ability to manipulate genes toward modeling and treating human disease. CRISPR-Cas9 gene regulation-modulating therapy is now in the clinic. CRISPR-Cas9 has dramatically simplified the process of modeling disease-associated mutations in pluripotent stem cells and their derivatives. More ambitiously, CRISPRa is beginning to be used to transform cellular identity through reprogramming and to correct diseases of insufficient gene expression. Such approaches are currently in their infancy, but given the pace of the CRISPR-Cas9 field, the ability to control cell fate may not be too far in the future.

In sum, the explosion of potent and precise techniques to edit genome sequence and epigenetic functionalization of this sequence is pushing us closer to true understanding of the complexities of development and the ability to precisely tweak gene regulation at will.

References

Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., Verdine, V., Cox, D. B. T., Kellner, M. J., Regev, A. et al. (2017). RNA targeting with CRISPR–Cas13. *Nature* 550, 280-284. doi:10.1038/nature24049

Adamson, B., Norman, T. M., Jost, M., Cho, M. Y., Nuñez, J. K., Chen, Y., Villalta, J. E., Gilbert, L. A., Horlbeck, M. A., Hein, M. Y. et al. (2016). A Multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* 167, 1867-1882.e21. doi:10.1016/j.cell.2016.11.048

Aguet, F., Brown, A. A., Castel, S. E., Davis, J. R., He, Y., Jo, B., Mohammadi, P., Park, Y., Parsana, P., Segrè, A. V. et al. (2017). Genetic effects on gene expression across human tissues. *Nature* 550, 204-213. doi:10.1038/nature24277

Allen, F., Crepaldi, L., Alsinet, C., Strong, A. J., Kleshchevnikov, V., De Angeli, P., Páleníková, P., Khodak, A., Kiselev, V., Kosicki, M. et al. (2019). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat. Biotechnol.* 37, 64-72. doi:10.1038/nbt.4317

Allis, C. D. and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 17, 487-500. doi:10.1038/nrg.2016.59

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T. et al. (2014). An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455-461. doi:10.1038/nature12787

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A. et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149-157. doi:10.1038/s41586-019-1711-4

Anzalone, A. V., Koblan, L. W. and Liu, D. R. (2020). Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38, 824-844. doi:10.1038/s41587-020-0561-9

Arnold, C. D., Gerlach, D., Stelzer, C., Boryń, Ł. M., Rath, M. and Stark, A. (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074-1077. doi:10.1126/science.1232542

Ashcroft, F. M. and Rorsman, P. (2012). Diabetes mellitus and the β cell: the last ten years. *Cell* 148, 1160-1171. doi:10.1016/j.cell.2012.02.010

Banerjee, B. and Sherwood, R. I. (2017). A CRISPR view of gene regulation. *Curr. Opin. Syst. Biol.* 1, 1-8. doi:10.1016/j.coisb.2016.12.016

Barash, Y., Calarco, J. A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B. J. and Frey, B. J. (2010). Deciphering the splicing code. *Nature* 465, 53-59. doi:10.1038/nature09000

Bauer, D. E., Kamran, S. C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M. C., Smith, E. C., Pinello, L. et al. (2013). An Erythroid Enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 342, 253-257. doi:10.1126/science.1242088

Bennett, C. F. (2019). Therapeutic antisense oligonucleotides are coming of age. *Annu. Rev. Med.* 70, 307-321. doi:10.1146/annurev-med-041217-010829

Black, J. B., Adler, A. F., Wang, H.-G., D'Ippolito, A. M., Hutchinson, H. A., Reddy, T. E., Pitt, G. S., Leong, K. W. and Gersbach, C. A. (2016). Targeted epigenetic remodeling of endogenous loci by CRISPR/Cas9-based transcriptional activators directly converts fibroblasts to neuronal cells. *Cell Stem Cell* 19, 406-414. doi:10.1016/j.stem.2016.07.001

Boyle, E. A., Li, Y. I. and Pritchard, J. K. (2017). An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169, 1177-1186. doi:10.1016/j.cell.2017.05.038

Braun, S. M. G., Kirkland, J. G., Chory, E. J., Husmann, D., Calarco, J. P. and Crabtree, G. R. (2017). Rapid and reversible epigenome editing by endogenous chromatin regulators. *Nat. Commun.* 8, 560. doi:10.1038/s41467-017-00644-y

Burgio, G. (2018). Redefining mouse transgenesis with CRISPR/Cas9 genome editing technology. *Genome Biol.* 19, 27. doi:10.1186/s13059-018-1409-1

Cacheiro, P., Haendel, M. A. and Smedley, D. and **International Mouse Phenotyping Consortium and the Monarch Initiative.** (2019). New models for human disease from the international mouse phenotyping consortium. *Mamm. Genome* 30, 143-150. doi:10.1007/s00335-019-09804-5

Cahan, P., Li, H., Morris, S. A., Lummertz da Rocha, E., Daley, G. Q. and Collins, J. J. (2014). CellNet: network biology applied to stem cell engineering. *Cell* 158, 903-915. doi:10.1016/j.cell.2014.07.020

Cano-Rodriguez, D., Gjaltema, R. A. F., Jilderda, L. J., Jellema, P., Dokter-Fokkens, J., Ruiters, M. H. J. and Rots, M. G. (2016). Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. *Nat. Commun.* 7, 12284. doi:10.1038/ncomms12284

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. et al. (2015). BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* 527, 192-197. doi:10.1038/nature15521

Canver, M. C., Lessard, S., Pinello, L., Wu, Y., Ilboudo, Y., Stern, E. N., Needleman, A. J., Galactéros, F., Brugnara, C., Kutlar, A. et al. (2017). Variant-aware saturating mutagenesis using multiple Cas9 nucleases identifies regulatory elements at trait-associated loci. *Nat. Genet.* 49, 625-634. doi:10.1038/ng.3793

Centonze, A., Lin, S., Tika, E., Sifrim, A., Fioramonti, M., Malfait, M., Song, Y., Wuidart, A., Van Herck, J., Dannau, A. et al. (2020). Heterotypic cell–cell communication regulates glandular stem cell multipotency. *Nature* 584, 608-613. doi:10.1038/s41586-020-2632-y

Chacón-Martínez, C. A., Koester, J. and Wickström, S. A. (2018). Signaling in the stem cell niche: regulating cell fate, function and plasticity. *Development* 145, dev165399. doi:10.1242/dev.165399

Charlesworth, C. T., Deshpande, P. S., Dever, D. P., Camarena, J., Lemgart, V. T., Cromer, M. K., Vakulskas, C. A., Collingwood, M. A., Zhang, L., Bode, N. M. et al. (2019). Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* 25, 249-254. doi:10.1038/s41591-018-0326-x

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., E. P. R., Iyer, Lin, S., Kiani, S., Guzman, C. D., Wiegand, D. J. et al. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* 12, 326-328. doi:10.1038/nmeth.3312

Chavez, A., Tuttle, M., Pruitt, B. W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S. J., Cecchi, R. J., Kowal, E. J. K., Buchthal, J. et al. (2016). Comparison of Cas9 activators in multiple species. *Nat. Methods* 13, 563-567. doi:10.1038/nmeth.3871

Cherry, A. B. C. and Daley, G. Q. (2012). Reprogramming cellular identity for regenerative medicine. *Cell* 148, 1110-1122. doi:10.1016/j.cell.2012.02.031

Chew, W. L., Tabebordbar, M., Cheng, J. K. W., Mali, P., Wu, E. Y., Ng, A. H. M., Zhu, K., Wagers, A. J. and Church, G. M. (2016). A multifunctional AAV–CRISPR–Cas9 and its host response. *Nat. Methods* 13, 868-874. doi:10.1038/nmeth.3993

Cohen, D. E. and Melton, D. (2011). Turning straw into gold: directing cell fate for regenerative medicine. *Nat. Rev. Genet.* 12, 243-252. doi:10.1038/nrg2938

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823. doi:10.1126/science.1231143

Cox, D. B. T., Gootenberg, J. S., Abudayyeh, O. O., Franklin, B., Kellner, M. J., Joung, J. and Zhang, F. (2017). RNA editing with CRISPR-Cas13. *Science* 358, 1019-1027. doi:10.1126/science.aaq0180

Davidson, E. H. (2010). Emerging properties of animal gene regulatory networks. *Nature* 468, 911-920. doi:10.1038/nature09645

Degner, J. F., Pai, A. A., Pique-Regi, R., Veyrieras, J.-B., Gaffney, D. J., Pickrell, J. K., De Leon, S., Michelini, K., Lewellen, N., Crawford, G. E. et al. (2012). DNaseI sensitivity QTLs are a major determinant of human expression variation. *Nature* 482, 390-394. doi:10.1038/nature10808

Diao, Y., Li, B., Meng, Z., Jung, I., Lee, A. Y., Dixon, J., Maliskova, L., Guan, K.-L., Shen, Y. and Ren, B. (2016). A new class of temporally phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening. *Genome Res.* **26**, 397-405. doi:10.1101/gr.197152.115

Diao, Y., Fang, R., Li, B., Meng, Z., Yu, J., Qiu, Y., Lin, K. C., Huang, H., Liu, T., Marina, R. J. et al. (2017). A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat. Methods* **14**, 629-635. doi:10.1038/nmeth.4264

Dickinson, M. E., Flenniken, A. M., Ji, X., Teboul, L., Wong, M. D., White, J. K., Meehan, T. F., Weninger, W. J., Westerberg, H., Adissu, H. et al. (2016). High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508-514. doi:10.1038/nature19356

Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., Marjanovic, N. D., Dionne, D., Burks, T., Raychowdhury, R. et al. (2016). Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853-1866.e17. doi:10.1016/j.cell.2016.11.038

Dolan, A. E., Hou, Z., Xiao, Y., Gramelspacher, M. J., Heo, J., Howden, S. E., Freddolino, P. L., Ke, A. and Zhang, Y. (2019). Introducing a spectrum of long-range genomic deletions in human embryonic stem cells using Type I CRISPR-Cas. *Mol. Cell* **74**, 936-950.e5. doi:10.1016/j.molcel.2019.03.014

Dunbar, C. E., High, K. A., Joung, J. K., Kohn, D. B., Ozawa, K. and Sadelain, M. (2018). Gene therapy comes of age. *Science* **359**, eaan4672. doi:10.1126/science.aan4672

Engreitz, J. M., Haines, J. E., Perez, E. M., Munson, G., Chen, J., Kane, M., McDonel, P. E., Guttman, M. and Lander, E. S. (2016). Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452-455. doi:10.1038/nature20149

Fahed, A. C., Wang, M., Homburger, J. R., Patel, A. P., Bick, A. G., Neben, C. L., Lai, C., Brockman, D., Philippakis, A., Ellinor, P. T. et al. (2020). Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. *Nat. Commun.* **11**, 3635. doi:10.1038/s41467-020-17374-3

Farzadfard, F., Perli, S. D. and Lu, T. K. (2013). Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synth. Biol.* **18**, 604-613. doi:10.1021/sb400081r

Fiaux, P. C., Chen, H. V., Chen, P. B., Chen, A. R. and McVicker, G. (2020). Discovering functional sequences with RELICS, an analysis method for CRISPR screens. *PLoS Comput. Biol.* **16**, e1008194. doi:10.1371/journal.pcbi.1008194

Fitz, J., Neumann, T., Steininger, M., Wiedemann, E.-M., Garcia, A. C., Athanasiadis, A., Schoeberl, U. E. and Pavri, R. (2020). Spt5-mediated enhancer transcription directly couples enhancer activation with physical promoter interaction. *Nat. Genet.* **52**, 505-515. doi:10.1038/s41588-020-0605-6

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R. et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and β-Thalassemia. *N. Engl. J. Med.* **384**, 252-260. doi:10.1056/NEJMoa2031054

Fu, X., Wu, X., Djekidel, M. N. and Zhang, Y. (2019). Myc and Dnmt1 impede the pluripotent to totipotent state transition in embryonic stem cells. *Nat. Cell Biol.* **21**, 835-844. doi:10.1038/s41556-019-0343-0

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. doi:10.1126/science.aag2445

Fulco, C. P., Nasser, J., Jones, T. R., Munson, G., Bergman, D. T., Subramanian, V., Grossman, S. R., Anyoha, R., Doughty, B. R., Patwardhan, T. A. et al. (2019). Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664-1669. doi:10.1038/s41588-019-0538-0

Gallagher, M. D. and Chen-Plotkin, A. S. (2018). The Post-GWAS Era: from association to function. *Am. J. Hum. Genet.* **102**, 717-730. doi:10.1016/j.ajhg.2018.04.002

Gao, X., Tsang, J. C., 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.* **10**, e155. doi:10.1093/nar/gku836

Gasperini, M., Hill, A. J., McFaline-Figueroa, J. L., Martin, B., Kim, S., Zhang, M. D., Jackson, D., Leith, A., Schreiber, J., Noble, W. S. et al. (2019). A genome-wide framework for mapping gene regulation via cellular genetic screens. *Cell* **176**, 377-390.e19. doi:10.1016/j.cell.2018.11.029

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I. and Liu, D. R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464-471. doi:10.1038/nature24644

Genga, R. M. J., Kernfeld, E. M., Parsi, K. M., Parsons, T. J., Ziller, M. J. and Maehr, R. (2019). Single-cell RNA-sequencing-based CRISPRi screening resolves molecular drivers of early human endoderm development. *Cell Rep.* **27**, 708-718.e10. doi:10.1016/j.celrep.2019.03.076

Giladi, A., Cohen, M., Medaglia, C., Baran, Y., Li, B., Zada, M., Bost, P., Blecher-Gonen, R., Salame, T.-M., Mayer, J. U. et al. (2020). Dissecting cellular crosstalk by sequencing physically interacting cells. *Nat. Biotechnol.* **38**, 629-637. doi:10.1038/s41587-020-0442-2

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. et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-451. doi:10.1016/j.cell.2013.06.044

Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., Guimaraes, C., Panning, B., Ploegh, H. L., Bassik, M. C. et al. (2014). Genome-Scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647-661. doi:10.1016/j.cell.2014.09.029

Gonzalez-Sandoval, A. and Gasser, S. M. (2016). On TADs and LADs: spatial control over gene expression. *Trends Genet.* **32**, 485-495. doi:10.1016/j.tig.2016.05.004

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D. U., Jung, I., Wu, H., Zhai, Y., Tang, Y. et al. (2015). CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900-910. doi:10.1016/j.cell.2015.07.038

Hammelman, J., Krismer, K., Banerjee, B., Gifford, D. K. and Sherwood, R. (2020). Identification of determinants of differential chromatin accessibility through a massively parallel genome-integrated reporter assay. *Genomics* **30**,1468-1480. doi:10.1101/gr.263228.120

Hart, T. and Moffat, J. (2016). BAGEL: a computational framework for identifying essential genes from pooled library screens. *BMC Bioinformatics* **17**, 164. doi:10.1186/s12859-016-1015-8

He, P., Williams, B. A., Trout, D., Marinov, G. K., Amrhein, H., Berghella, L., Goh, S.-T., Plajzer-Frick, I., Afzal, V., Pennacchio, L. A. et al. (2020). The changing mouse embryo transcriptome at whole tissue and single-cell resolution. *Nature* **583**, 760-767. doi:10.1038/s41586-020-2536-x

Henninger, J. E., Oksuz, O., Shrinivas, K., Sagi, I., LeRoy, G., Zheng, M. M., Andrews, J. O., Zamudio, A. V., Lazaris, C., Hannett, N. M. et al. (2021). RNA-mediated feedback control of transcriptional condensates. *Cell* **184**, 207-225.e24. doi:10.1016/j.cell.2020.11.030

Henrique, D. and Schweisguth, F. (2019). Mechanisms of Notch signaling: a simple logic deployed in time and space. *Development* **146**, dev172148. doi:10.1242/dev.172148

Henriques, T., Scruggs, B. S., Inouye, M. O., Muse, G. W., Williams, L. H., Burkholder, A. B., Lavender, C. A., Fargo, D. C. and Adelman, K. (2018). Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev.* **32**, 26-41. doi:10.1101/gad.309351.117

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. doi:10.1038/nbt.3199

Horlbeck, M. A., Gilbert, L. A., Villalta, J. E., Adamson, B., Pak, R. A., Chen, Y., Fields, A. P., Park, C. Y., Corn, J. E., Kampmann, M. et al. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760. doi:10.7554/eLife.19760

Hsu, J. Y., Fulco, C. P., Cole, M. A., Canver, M. C., Pellin, D., Sher, F., Farouni, R., Clement, K., Guo, J. A., Biasco, L. et al. (2018). CRISPR-SURF: discovering regulatory elements by deconvolution of CRISPR tiling screen data. *Nat. Methods* **15**, 992-993. doi:10.1038/s41592-018-0225-6

Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S. H. and Elledge, S. J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev.* **23**, 837-848. doi:10.1101/gad.1769609

Inoue, F. and Ahituv, N. (2015). Decoding enhancers using massively parallel reporter assays. *Genomics* **106**, 159-164. doi:10.1016/j.ygeno.2015.06.005

Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., Kosmicki, J. A., Arbelaez, J., Cui, W., Schwartz, G. B. et al. (2019). Predicting splicing from primary sequence with deep learning. *Cell* **176**, 535-548.e24. doi:10.1016/j.cell.2018.12.015

Jaitin, D. A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., Salame, T. M., Tanay, A., van Oudenaarden, A. and Amit, I. (2016). Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-Seq. *Cell* **167**, 1883-1896.e15. doi:10.1016/j.cell.2016.11.039

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. and Doudna, J. (2013). RNA-programmed genome editing in human cells. *eLife* **2**, e00471. doi:10.7554/eLife.00471

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. doi:10.1038/nmeth.3325

Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., Natarajan, P., Lander, E. S., Lubitz, S. A., Ellinor, P. T. et al. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219-1224. doi:10.1038/s41588-018-0183-z

Khoury, A., Achinger-Kawecka, J., Bert, S. A., Smith, G. C., French, H. J., Luu, P.-L., Peters, T. J., Du, Q., Parry, A. J., Valdes-Mora, F. et al. (2020). Constitutively bound CTCF sites maintain 3D chromatin architecture and long-range epigenetically regulated domains. *Nat. Commun.* **11**, 54. doi:10.1038/s41467-019-13753-7

Kilpinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., Bala, S., Bensaddek, D., Casale, F. P., Culley, O. J. et al. (2017). Common genetic

variation drives molecular heterogeneity in human iPSCs. *Nature* **546**, 370-375. doi:10.1038/nature22403

Kim, T.-K., Hemberg, M., Gray, J. M., Costa, A. M., Bear, D. M., Wu, J., Harmin, D. A., Laptewicz, M., Barbara-Haley, K., Kuersten, S. et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182-187. doi:10.1038/nature09033

Kim, J. H., Rege, M., Valeri, J., Dunagin, M. C., Metzger, A., Titus, K. R., Gilgenast, T. G., Gong, W., Beagan, J. A., Raj, A. et al. (2019a). LADL: light-activated dynamic looping for endogenous gene expression control. *Nat. Methods* **16**, 633-639. doi:10.1038/s41592-019-0436-5

Kim, H. K., Kim, Y., Lee, S., Min, S., Bae, J. Y., Choi, J. W., Park, J., Jung, D., Yoon, S. and Kim, H. H. (2019b). SpCas9 activity prediction by DeepSpCas9, a deep learning–based model with high generalization performance. *Sci. Adv.* **5**, eaax9249. doi:10.1126/sciadv.aax9249

Kim, N., Kim, H. K., Lee, S., Seo, J. H., Choi, J. W., Park, J., Min, S., Yoon, S., Cho, S.-R. and Kim, H. H. (2020). Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nat. Biotechnol.* **38**, 1328-1336. doi:10.1038/s41587-020-0537-9

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M. and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310-315. doi:10.1038/ng.2892

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. doi:10.1038/nbt.3853

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., Gonzales, A. P. W., Li, Z., Peterson, R. T., Yeh, J.-R. J. et al. (2015a). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481-485. doi:10.1038/nature14592

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Topkar, V. V., Zheng, Z. and Joung, J. K. (2015b). Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* **33**, 1293-1298. doi:10.1038/nbt.3404

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z. and Joung, J. K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495. doi:10.1038/nature16526

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424. doi:10.1038/nature17946

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., Hsu, P. D., Habib, N., Gootenberg, J. S., Nishimasu, H. et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583-588. doi:10.1038/nature14136

Kopp, F. and Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell* **172**, 393-407. doi:10.1016/j.cell.2018.01.011

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. doi:10.1038/nbt.3450

Korkmaz, G., Manber, Z., Lopes, R., Prekovic, S., Schuurman, K., Kim, Y., Teunissen, H., Flach, K., de Wit, E. et al. (2019). A CRISPR-Cas9 screen identifies essential CTCF anchor sites for estrogen receptor-driven breast cancer cell proliferation. *Nucleic Acids Res.* **47**, 9557-9572. doi:10.1093/nar/gkz675

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R. and Weirauch, M. T. (2018). The human transcription factors. *Cell* **172**, 650-665. doi:10.1016/j.cell.2018.01.029

Lareau, C. A., Duarte, F. M., Chew, J. G., Kartha, V. K., Burkett, Z. D., Kohlway, A. S., Pokholok, D., Aryee, M. J., Steemers, F. J., Lebofsky, R. et al. (2019). Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat. Biotechnol.* **37**, 916-924. doi:10.1038/s41587-019-0147-6

Li, W., Xu, H., Xiao, T., Cong, L., Love, M. I., Zhang, F., Irizarry, R. A., Liu, J. S., Brown, M. and Liu, X. S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554. doi:10.1186/s13059-014-0554-4

Li, M., Yu, J. S. L., Tilgner, K., Ong, S. H., Koike-Yusa, H. and Yusa, K. (2018). Genome-wide CRISPR-KO Screen Uncovers mTORC1-mediated Gsk3 regulation in naive pluripotency maintenance and dissolution. *Cell Rep.* **24**, 489-502. doi:10.1016/j.celrep.2018.06.027

Li, Q. V., Dixon, G., Verma, N., Rosen, B. P., Gordillo, M., Luo, R., Xu, C., Wang, Q., Soh, C.-L., Yang, D. et al. (2019). Genome-scale screens identify JNK–JUN signaling as a barrier for pluripotency exit and endoderm differentiation. *Nat. Genet.* **51**, 999-1010. doi:10.1038/s41588-019-0408-9

Liu, Y., Yu, C., Daley, T. P., Wang, F., Cao, W. S., Bhate, S., Lin, X., Still, C., Liu, H., Zhao, D. et al. (2018). CRISPR activation screens systematically identify factors that drive neuronal fate and reprogramming. *Cell Stem Cell* **23**, 758-771.e8. doi:10.1016/j.stem.2018.09.003

Luecken, M. D. and Theis, F. J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **19**, e8746. doi:10.15252/msb.20188746

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M. et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202-1214. doi:10.1016/j.cell.2015.05.002

Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H. and Joung, J. K. (2013). CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977-979. doi:10.1038/nmeth.2598

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. doi:10.1126/science.1232033

Manghwar, H., Lindsey, K., Zhang, X. and Jin, S. (2019). CRISPR/Cas system: recent advances and future prospects for genome editing. *Trends Plant Sci.* **24**, 1102-1125. doi:10.1016/j.tplants.2019.09.006

Matharu, N., Rattanasopha, S., Tamura, S., Maliskova, L., Wang, Y., Bernard, A., Hardin, A., Eckalbar, W. L., Vaisse, C. and Ahituv, N. (2019). CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* **363**, eaau0629. doi:10.1126/science.aau0629

McVicker, G., van de Geijn, B., Degner, J. F., Cain, C. E., Banovich, N. E., Raj, A., Lewellen, N., Myrthil, M., Gilad, Y. and Pritchard, J. K. (2013). Identification of genetic variants that affect histone modifications in human cells. *Science* **342**, 747-749. doi:10.1126/science.1242429

Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C. G., Jr., Kinney, J. B. et al. (2012). Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* **30**, 271-277. doi:10.1038/nbt.2137

Mihelčić, M., Šmuc, T. and Supek, F. (2019). Patterns of diverse gene functions in genomic neighborhoods predict gene function and phenotype. *Sci. Rep.* **9**, 19537. doi:10.1038/s41598-019-55984-0

Morgan, S. L., Mariano, N. C., Bermudez, A., Arruda, N. L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.-F. et al. (2017). Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun.* **8**, 15993. doi:10.1038/ncomms15993

Moris, N., Anlas, K., van den Brink, S. C., Alemany, A., Schröder, J., Ghimire, S., Balayo, T., van Oudenaarden, A. and Martinez Arias, A. (2020). An in vitro model of early anteroposterior organization during human development. *Nature* **582**, 410-415. doi:10.1038/s41586-020-2383-9

Morisaka, H., Yoshimi, K., Okuzaki, Y., Gee, P., Kunihiro, Y., Sonpho, E., Xu, H., Sasakawa, N., Naito, Y., Nakada, S. et al. (2019). CRISPR-Cas3 induces broad and unidirectional genome editing in human cells. *Nat. Commun.* **10**, 5302. doi:10.1038/s41467-019-13226-x

Musunuru, K. (2013). Genome editing of human pluripotent stem cells to generate human cellular disease models. *Dis. Model. Mech.* **6**, 896-904. doi:10.1242/dmm.012054

Musunuru, K., Sheikh, F., Gupta, R. M., Houser, S. R., Maher, K. O., Milan, D. J., Terzic, A., Wu, J. C., American Heart Association Council on Functional Genomics and Translational Biology; Council on Cardiovascular Disease in the Young. et al. (2018). Induced pluripotent stem cells for cardiovascular disease modeling and precision medicine: a scientific statement from the american heart association. *Circ. Genomic Precis. Med.* **11**, e000043. doi:10.1161/HCG.0000000000000043

Nirenberg, M., Caskey, T., Marshall, R., Brimacombe, R., Kellogg, D., Doctor, B., Hatfield, D., Levin, J., Rottman, F., Pestka, S. et al. (1966). The RNA code and protein synthesis. *Cold Spring Harb. Symp. Quant. Biol.* **31**, 11-24. doi:10.1101/SQB.1966.031.01.008

Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O. O., Gootenberg, J. S., Mori, H. et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361**, 1259-1262. doi:10.1126/science.aas9129

Nishizaki, S. S. and Boyle, A. P. (2017). Mining the unknown: assigning function to noncoding single nucleotide polymorphisms. *Trends Genet.* **33**, 34-45. doi:10.1016/j.tig.2016.10.008

O'Connell, M. R. (2019). Molecular mechanisms of RNA targeting by Cas13-containing Type VI CRISPR-Cas systems. *J. Mol. Biol.* **431**, 66-87. doi:10.1016/j.jmb.2018.06.029

O'Geen, H., Bates, S. L., Carter, S. S., Nisson, K. A., Halmai, J., Fink, K. D., Rhie, S. K., Farnham, P. J. and Segal, D. J. (2019). Ezh2-dCas9 and KRAB-dCas9 enable engineering of epigenetic memory in a context-dependent manner. *Epigenetics Chromatin* **12**, 26. doi:10.1186/s13072-019-0275-8

Pagliuca, F. W., Millman, J. R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J. H., Peterson, Q. P., Greiner, D. and Melton, D. A. (2014). Generation of functional human pancreatic β cells in vitro. *Cell* **159**, 428-439. doi:10.1016/j.cell.2014.09.040

Parnas, O., Jovanovic, M., Eisenhaure, T. M., Herbst, R. H., Dixit, A., Ye, C. J., Przybylski, D., Platt, R. J., Tirosh, I., Sanjana, N. E. et al. (2015). A Genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675-686. doi:10.1016/j.cell.2015.06.059

Partridge, E. C., Chhetri, S. B., Prokop, J. W., Ramaker, R. C., Jansen, C. S., Goh, S.-T., Mackiewicz, M., Newberry, K. M., Brandsmeier, L. A., Meadows, S. K. et al. (2020). Occupancy maps of 208 chromatin-associated proteins in one human cell type. *Nature* **583**, 720-728. doi:10.1038/s41586-020-2023-4

Pausch, P., Al-Shayeb, B., Bison-Rapp, E., Tsuchida, C. A., Li, Z., Cress, B. F., Knott, G. J., Jacobsen, S. E., Banfield, J. F. and Doudna, J. A. (2020). CRISPR-CasΦ from huge phages is a hypercompact genome editor. *Science* **369**, 333-337.

Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., Thakore, P. I., Glass, K. A., Ousterout, D. G., Leong, K. W. et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973-976. doi:10.1038/nmeth.2600

Pickar-Oliver, A. and Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490-507. doi:10.1038/s41580-019-0131-5

Popp, B., Krumbiegel, M., Grosch, J., Sommer, A., Uebe, S., Kohl, Z., Plötz, S., Farrell, M., Trautmann, U., Kraus, C. et al. (2018). Need for high-resolution genetic analysis in iPSC: results and lessons from the ForIPS consortium. *Sci. Rep.* **8**, 17201. doi:10.1038/s41598-018-35506-0

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* **28**, 1173-1183. doi:10.1016/j.cell.2013.02.022

Qian, H., Kang, X., Hu, J., Zhang, D., Liang, Z., Meng, F., Zhang, X., Xue, Y., Maimon, R., Dowdy, S. F. et al. (2020). Reversing a model of Parkinson's disease with in situ converted nigral neurons. *Nature* **582**, 550-556. doi:10.1038/s41586-020-2388-4

Rackham, O. J. L., Firas, J., Fang, H., Oates, M. E., Holmes, M. L., Knaupp, A. S., Suzuki, H., Nefzger, C. M., Daub, C. O., Shin, J. W. et al. (2016). A predictive computational framework for direct reprogramming between human cell types. *Nat. Genet.* **48**, 331-335. doi:10.1038/ng.3487

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. doi:10.1038/nbt.3468

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M. et al. (2017). Science forum: the human cell atlas. *eLife* **6**, e27041. doi:10.7554/eLife.27041

Reilly, S. K. and Noonan, J. P. (2016). Evolution of gene regulation in humans. *Annu. Rev. Genomics Hum. Genet.* **17**, 45-67. doi:10.1146/annurev-genom-090314-045935

Reiter, F., Wienerroither, S. and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. *Curr. Opin. Genet. Dev.* **43**, 73-81. doi:10.1016/j.gde.2016.12.007

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J. et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-330. doi:10.1038/nature14248

Rowe, R. G. and Daley, G. Q. (2019). Induced pluripotent stem cells in disease modelling and drug discovery. *Nat. Rev. Genet.* **20**, 377-388. doi:10.1038/s41576-019-0100-z

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. doi:10.1126/science.aaf7613

Shahbazi, M. N. (2020). Mechanisms of human embryo development: from cell fate to tissue shape and back. *Development* **147**, dev190629. doi:10.1242/dev.190629

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelson, T. S., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G. et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science.* **343**, 84-87. doi:10.1126/science.1247005

Shechner, D. M., Hacisuleyman, E., Younger, S. T. and Rinn, J. L. (2015). Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat. Methods* **12**, 664-670. doi:10.1038/nmeth.3433

Shen, M. W., Arbab, M., Hsu, J. Y., Worstell, D., Culbertson, S. J., Krabbe, O., Cassa, C. A., Liu, D. R., Gifford, D. K. and Sherwood, R. I. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646-651. doi:10.1038/s41586-018-0686-x

Sherwood, R. I., Hashimoto, T., O'Donnell, C. W., Lewis, S., Barkal, A. A., van Hoff, J. P., Karun, V., Jaakkola, T. and Gifford, D. K. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**, 171-178. doi:10.1038/nbt.2798

Shi, J., Wang, E., Milazzo, J. P., Wang, Z., Kinney, J. B. and Vakoc, C. R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 661-667. doi:10.1038/nbt.3235

Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X. and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84-88. doi:10.1126/science.aad5227

Smargon, A. A., Shi, Y. J. and Yeo, G. W. (2020). RNA-targeting CRISPR systems from metagenomic discovery to transcriptomic engineering. *Nat. Cell Biol.* **22**, 143-150. doi:10.1038/s41556-019-0454-7

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. doi:10.1038/ng.2713

Smith, A. J. P., Deloukas, P. and Munroe, P. B. (2018). Emerging applications of genome-editing technology to examine functionality of GWAS-associated variants for complex traits. *Physiol. Genomics* **50**, 510-522. doi:10.1152/physiolgenomics.00028.2018

Spitz, F. and Furlong, E. E. M. (2012). Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613-626. doi:10.1038/nrg3207

Stergachis, A. B., Debo, B. M., Haugen, E., Churchman, L. S. and Stamatoyannopoulos, J. A. (2020). Single-molecule regulatory architectures captured by chromatin fiber sequencing. *Science* **368**, 1449-1454. doi:10.1126/science.aaz1646

Stunnenberg, H. G., Hirst, M., Abrignani, S., Adams, D., de Almeida, M., Altucci, L., Amin, V., Amit, I., Antonarakis, S. E., Aparicio, S. et al. (2016). The international human epigenome consortium: a blueprint for scientific collaboration and discovery. *Cell* **167**, 1145-1149. doi:10.1016/j.cell.2016.11.007

Szczesnik, T., Chu, L., Ho, J. W. K. and Sherwood, R. (2020). A high-throughput genome-integrated assay reveals spatial dependencies governing Tcf7l2 binding. *Cell Syst.* **11**, 315-327.e5. doi:10.1016/j.cels.2020.08.004

Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676. doi:10.1016/j.cell.2006.07.024

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G. and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **20**, 467-484. doi:10.1038/s41576-019-0127-1

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. and Vale, R. D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635-646. doi:10.1016/j.cell.2014.09.039

Tarjan, D. R., Flavahan, W. A. and Bernstein, B. E. (2019). Epigenome editing strategies for the functional annotation of CTCF insulators. *Nat. Commun.* **10**, 4258. doi:10.1038/s41467-019-12166-w

Tewhey, R., Kotliar, D., Park, D. S., Liu, B., Winnicki, S., Reilly, S. K., Andersen, K. G., Mikkelsen, T. S., Lander, E. S., Schaffner, S. F. et al. (2016). Direct identification of hundreds of expression-modulating variants using a multiplexed reporter assay. *Cell* **165**, 1519-1529. doi:10.1016/j.cell.2016.04.027

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. doi:10.1038/nmeth.3630

The ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74. doi:10.1038/nature11247

Tickle, C. and Urrutia, A. O. (2017). Perspectives on the history of evo-devo and the contemporary research landscape in the genomics era. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20150473. doi:10.1098/rstb.2015.0473

Töhönen, V., Katayama, S., Vesterlund, L., Jouhilahti, E.-M., Sheikhi, M., Madissoon, E., Filippini-Cattaneo, G., Jaconi, M., Johnsson, A., Bürglin, T. R. et al. (2015). Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nat. Commun.* **6**, 8207. doi:10.1038/ncomms9207

van Arensbergen, J., Pagie, L., FitzPatrick, V. D., de Haas, M., Baltissen, M. P., Comoglio, F., van der Weide, R. H., Teunissen, H., Võsa, U., Franke, L. et al. (2019). High-throughput identification of human SNPs affecting regulatory element activity. *Nat. Genet.* **51**, 1160-1169. doi:10.1038/s41588-019-0455-2

van Overbeek, M., Capurso, D., Carter, M. M., Thompson, M. S., Frias, E., Russ, C., Reece-Hoyes, J. S., Nye, C., Gradia, S., Vidal, B. et al. (2016). DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. *Mol. Cell* **63**, 633-646. doi:10.1016/j.molcel.2016.06.037

Veres, A., Faust, A. L., Bushnell, H. L., Engquist, E. N., Kenty, J. H.-R., Harb, G., Poh, Y.-C., Sintov, E., Gürtler, M., Pagliuca, F. W. et al. (2019). Charting cellular identity during human in vitro β-cell differentiation. *Nature* **569**, 368-373. doi:10.1038/s41586-019-1168-5

Vierbuchen, T., Ostermeier, A., Pang, Z. P., Kokubu, Y., Südhof, T. C. and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035-1041. doi:10.1038/nature08797

Volpato, V., Smith, J., Sandor, C., Ried, J. S., Baud, A., Handel, A., Newey, S. E., Wessely, F., Attar, M., Whiteley, E. et al. (2018). Reproducibility of molecular phenotypes after long-term differentiation to human iPSC-derived neurons: a multi-site omics study. *Stem Cell Rep.* **11**, 897-911. doi:10.1016/j.stemcr.2018.08.013

Walton, R. T., Christie, K. A., Whittaker, M. N. and Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290-296. doi:10.1126/science.aba8853

Wang, D., Mou, H., Li, S., Li, Y., Hough, S., Tran, K., Li, J., Yin, H., Anderson, D. G., Sontheimer, E. J. et al. (2015). Adenovirus-mediated somatic genome

editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses. *Hum. Gene Ther.* **26**, 432-442. doi:10.1089/hum.2015.087

Wang, H., Xu, X., Nguyen, C. M., Liu, Y., Gao, Y., Lin, X., Daley, T., Kipniss, N. H., La Russa, M. and Qi, L. S. (2018). CRISPR-mediated programmable 3D genome positioning and nuclear organization. *Cell* **175**, 1405-1417.e14. doi:10.1016/j.cell.2018.09.013

Wang, D., Tai, P. W. L. and Gao, G. (2019a). Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* **18**, 358-378. doi:10.1038/s41573-019-0012-9

Wang, H., Nakamura, M., Abbott, T. R., Zhao, D., Luo, K., Yu, C., Nguyen, C. M., Lo, A., Daley, T. P., La Russa, M. et al. (2019b). CRISPR-mediated live imaging of genome editing and transcription. *Science* **365**, 1301-1305. doi:10.1126/science.aax7852

Wang, T., Wei, J. J., Sabatini, D. M. and Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **3**, 80-84. doi:10.1126/science.1246981

Weltner, J., Balboa, D., Katayama, S., Bespalov, M., Krjutškov, K., Jouhilahti, E.-M., Trokovic, R., Kere, J. and Otonkoski, T. (2018). Human pluripotent reprogramming with CRISPR activators. *Nat. Commun.* **9**, 2643. doi:10.1038/s41467-018-05067-x

Wilbrey-Clark, A., Roberts, K. and Teichmann, S. A. (2020). Cell Atlas technologies and insights into tissue architecture. *Biochem. J.* **477**, 1427-1442. doi:10.1042/BCJ20190341

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.e5. doi:10.1016/j.molcel.2017.03.007

Xu, J., Du, Y. and Deng, H. (2015). Direct lineage reprogramming: strategies, mechanisms, and applications. *Cell Stem Cell* **16**, 119-134. doi:10.1016/j.stem.2015.01.013

Xu, J., Zhou, C., Foo, K. S., Yang, R., Xiao, Y., Bylund, K., Sahara, M. and Chien, K. R. (2020). Genome-wide CRISPR screen identifies *ZIC2* as an essential gene that controls the cell fate of early mesodermal precursors to human heart progenitors. *Stem Cells* **38**, 741-755. doi:10.1002/stem.3168

Yang, L.-Z., Wang, Y., Li, S.-Q., Yao, R.-W., Luan, P.-F., Wu, H., Carmichael, G. G. and Chen, L.-L. (2019). Dynamic imaging of RNA in living cells by CRISPR-Cas13 systems. *Mol. Cell* **76**, 981-997.e7. doi:10.1016/j.molcel.2019.10.024

Yeo, G. H. T., Lin, L., Qi, C. Y., Cha, M., Gifford, D. K. and Sherwood, R. I. (2020). A multiplexed barcodelet single-cell RNA-seq approach Elucidates combinatorial signaling pathways that drive ESC differentiation. *Cell Stem Cell* **26**, 938-950.E6. doi:10.1016/j.stem.2020.04.020

Young, R. S., Kumar, Y., Bickmore, W. A. and Taylor, M. S. (2017). Bidirectional transcription initiation marks accessible chromatin and is not specific to enhancers. *Genome Biol.* **18**, 242. doi:10.1186/s13059-017-1379-8

Zabidi, M. A., Arnold, C. D., Schernhuber, K., Pagani, M., Rath, M., Frank, O. and Stark, A. (2015). Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**, 556-559. doi:10.1038/nature13994

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A. et al. (2015). Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. *Cell* **163**, 759-771. doi:10.1016/j.cell.2015.09.038

Zhang, X., Li, T., Liu, F., Chen, Y., Yao, J., Li, Z., Huang, Y. and Wang, J. (2019a). Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems. *Mol. Cell* **73**, 130-142.e5. doi:10.1016/j.molcel.2018.10.020

Zhang, Z., Lee, J.-H., Ruan, H., Ye, Y., Krakowiak, J., Hu, Q., Xiang, Y., Gong, J., Zhou, B., Wang, L. et al. (2019b). Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. *Nat. Commun.* **10**, 4562. doi:10.1038/s41467-019-12543-5

Zhou, H., Su, J., Hu, X., Zhou, C., Li, H., Chen, Z., Xiao, Q., Wang, B., Wu, W., Sun, Y. et al. (2020). Glia-to-neuron conversion by CRISPR-CasRx alleviates symptoms of neurological disease in mice. *Cell* **181**, 590-603.e16. doi:10.1016/j.cell.2020.03.024

Zhu, Z., Li, Q. V., Lee, K., Rosen, B. P., González, F., Soh, C.-L. and Huangfu, D. (2016). Genome editing of lineage determinants in human pluripotent stem cells reveals mechanisms of pancreatic development and diabetes. *Cell Stem Cell* **18**, 755-768. doi:10.1016/j.stem.2016.03.015

DEVELOPMENT