Enabling functional genomics with genome engineering

Isaac B. Hilton¹,² and Charles A. Gersbach¹,²,³

¹Department of Biomedical Engineering, Duke University, Durham, North Carolina 27708, USA; ²Center for Genomic and Computational Biology, Duke University, Durham, North Carolina 27708, USA; ³Department of Orthopaedic Surgery, Duke University Medical Center, Durham, North Carolina 27710, USA

Advances in genome engineering technologies have made the precise control over genome sequence and regulation possible across a variety of disciplines. These tools can expand our understanding of fundamental biological processes and create new opportunities for therapeutic designs. The rapid evolution of these methods has also catalyzed a new era of genomics that includes multiple approaches to functionally characterize and manipulate the regulation of genomic information. Here, we review the recent advances of the most widely adopted genome engineering platforms and their application to functional genomics. This includes engineered zinc finger proteins, TALEs/TALENs, and the CRISPR/Cas9 system as nucleases for genome editing, transcription factors for epigenome editing, and other emerging applications. We also present current and potential future applications of these tools, as well as their current limitations and areas for future advances.

Genomic research has the potential to dramatically improve medicine, agriculture, biotechnology, and our fundamental understanding of living systems. Recent advances have generated extensive annotation of genomic and epigenomic regulatory modules within chromatin (Bernstein et al. 2010; The ENCODE Project Consortium 2012; Roadmap Epigenomics Consortium et al. 2015), as well as an understanding of genomic topological architecture (Dekker et al. 2013; Pombo and Dillon 2015). However, the roles of these numerous genes, regulatory elements, epigenetic marks, and topological domains in determining overall cell function remain incompletely understood. The recent development of genome engineering technologies has enabled precise interrogation of the function of these genomic features and their causal role in gene regulation. Additionally, these tools are facilitating the translation of this genomic information into tangible benefits for biotechnology, agriculture, and human therapeutics. In this Perspective, we discuss the recent advances to the most commonly used genome engineering technologies, including synthetic zinc finger (ZF) proteins, transcription activator-like effectors (TALEs), and CRISPR/Cas9 targeting systems, and their application in a new era of functional genomics.

Genome engineering technologies

Cys2-His2 ZF domains are naturally occurring protein motifs which typically recognize three base pairs within the major groove of DNA (Pavletich and Pabo 1991; Wolfe et al. 2000). These modular ZF domains can be arrayed such that synthetic ZF DNA-binding proteins (DBPs) target a specific series of DNA triplets at unique genomic addresses (Fig. 1A; Liu et al. 1997; Gersbach et al. 2014). TALE proteins are components of plant pathogens that bind host DNA to facilitate virulence (Kay et al. 2007; Romer et al. 2007). TALEs consist of repeated DNA-binding domains containing repeat variable diresidues (RVDs), each of which recognizes a single nucleotide in target DNA (Boch et al. 2009; Moscou and Bogdanove 2009). Similar to ZFs, individual TALE RVDs can be linked in series to localize TALEs to target loci (Fig. 1B; Christian et al. 2010; Morbitzer et al. 2010; Cermak et al. 2011; Miller et al. 2011; Zhang et al. 2011). Clustered regularly interspaced short palindromic repeat (CRISPR) arrays and CRISPR-associated (Cas) proteins are components of bacterial and archaeal adaptive immune systems (Barrangou et al. 2007; Makarova et al. 2011). Unlike ZFs and TALEs, in which protein moieties dictate DNA recognition, CRISPR/Cas systems utilize RNA-mediated Watson-Crick bonding for recognition of nucleic acids.

Prokaryotes harboring type II CRISPR/Cas systems transcribe CRISPR-RNAs (crRNAs) that hybridize with trans-activating crRNAs (tracrRNAs) that complex with the Cas9 nuclease (Brouns et al. 2008; Deltcheva et al. 2011; Jinek et al. 2012; Doudna and Charpentier 2014). A single crRNA-tracrRNA chimeric, known as a guide RNA (gRNA), can be designed for simplified use in engineered systems (Jinek et al. 2012). The gRNA binds to and directs the Cas9 protein to DNA through regions of crRNA complementarity (termed “protospacer” sequences). A stringent prerequisite to protospacer hybridization is the presence of a protospacer adjacent motif (PAM) in the target DNA, which flanks the region of protospacer complementarity (Fig. 1C; Mojica et al. 2009; Anders et al. 2014; Sternberg et al. 2014; Kleinstiver et al. 2015). Interactions between the PAM-proximal “seed” nucleotides in the target site and the complementary gRNA sequence are also critical drivers of Cas9 targeting. The orthogonality of various prokaryotic Cas9 proteins with differing PAM requirements can be exploited for multiplex genome engineering efforts (Esveldt et al. 2013; Ran et al. 2015).

Engineered targeting of eukaryotic genomes with ZFs, TALEs, and type II CRISPR/Cas systems has established these technologies as useful resources for present and future genome engineering endeavors. Although each of these systems has been successfully incorporated into diverse genome engineering strategies, they each have unique benefits and limitations that depend upon the particular application (Gaj et al. 2013; Carroll 2014). Other genome engineering...
Functional genomics with genome engineering

engineering technologies, such as meganucleases and their fusion to TALE proteins (Boisse et al. 2014; Stodhard 2014) have also been successfully applied to gene editing in eukaryotes, but have been less widely adopted in the context of functional genomics due to the complexity involved in engineering meganucleases targeted to new sequences.

Editing genome sequences with programmable nucleases

Gene targeting based on homologous recombination can introduce exogenous DNA at genomic loci (Smithies et al. 1985; Thomas et al. 1986). The efficiency of this method is dramatically enhanced in the presence of double-strand breaks (DSBs) (Rouet et al. 1994). Cells generally use two distinct pathways to resolve DSBs (Chapman et al. 2012): homology-directed repair (HDR) and nonhomologous end joining (NHEJ) (Fig. 1D). DSB resolution through NHEJ occurs by direct ligation of DSB ends or through microhomology on DSB termini (Lieber 2010). This error-prone process results in small insertions or deletions (indels) at endogenous loci. The generation of two DSBs flanking a genomic region can also lead to NHEJ-mediated chromosomal deletions (Lee et al. 2010; Carlson et al. 2012; Kim et al. 2013; Essletzbichler et al. 2014) and inversions (Carlson et al. 2012; Lee et al. 2012; Xiao et al. 2013). Programmable nucleases can also be used to create NHEJ-mediated translocations in vivo (Brunet et al. 2009; Maddalo et al. 2014). HDR uses regions of homologous DNA on sister chromatids or exogenous DNA to repair DSBs. In contrast to NHEJ, HDR is a high-fidelity process leading to largely error-free correction at DSB sites. HDR can lead to specified incorporation of sequences that are rationally designed into donor DNA templates. In many instances, this is advantageous over NHEJ, in which DSB resolution is unpredictable. However, NHEJ is active
throughout the cell cycle and is the predominant DSB repair mechanism, whereas HDR occurs less frequently and is significantly down-regulated outside of S and G2 cell cycle phases. Therefore recent efforts have focused on promoting HDR by inhibition of NHEJ events to afford more precise genome editing (Chu et al. 2015; Maruyama et al. 2015).

Because NHEJ and HDR can be used to incorporate specific sequence changes into genomes, the capability to induce DSBs at target loci holds great potential for genome engineering (Carroll 2014; Kim and Kim 2014). Although ZFs and TALEs do not intrinsically cleave target DNA, they can be directly fused to the catalytic domain of the type II restriction endonuclease FokI (Li et al. 1992) to create ZF and TALE nuclease (ZFNs and TALENs). Engineering new DNA-binding specificities into the ZFs and TALEs allows programmed DSB induction (Kim et al. 1996; Christian et al. 2010). FokI domains must dimerize to cleave target DNA (Bitinaite et al. 1998; Vananee et al. 2003), necessitating the engineering of two ZFN or TALE monomers for each DSB. Modifications to the FokI domain have also been created to increase nuclease activity and specificity (Miller et al. 2007; Szczepak et al. 2007; Guo et al. 2010; Doyon et al. 2011b). Moreover, fusion of the FokI domain to the nuclease-inactivated Cas9 protein (dCas9) has also been used to increase the specificity of the CRISPR gene editing system (Guilinger et al. 2014b; Tsai et al. 2014). Optimizing the length of the ZF or TALE array can also increase nuclease activity and specificity (Bhakta et al. 2013; Guilinger et al. 2014a).

As an alternative to introducing DSBs and inducing DNA repair pathways, catalytic domains of site-specific recombinases can also be fused to synthetic ZFs (Akopian et al. 2003) and TALEs (Merriner et al. 2012) to excise genomic DNA segments or integrate exogenous DNA at targeted genomic sites (Gordley et al. 2011). CRISPR/Cas-based recombinase fusions have not yet been reported but may also prove useful. Additionally, transposases catalyze the rearrangement of endogenous elements and have been used for artificial genome manipulation (Ivics et al. 2009). Targeting of transposase activity with ZF, TALE, or CRISPR/Cas scaffolds is also an active area of research (Yant et al. 2007; Voigt et al. 2012; Galvan et al. 2014). Importantly, targeted recombination or transposition may reduce cellular toxicity relative to the introduction of DSBs and their subsequent resolution through NHEJ or HDR. These alternate mechanisms to genome engineering may also increase efficiencies by decoupling editing from endogenous DNA repair mechanisms.

Synthetic regulation of transcription

In addition to the editing of DNA sequences, these genome engineering technologies can be used to manipulate endogenous gene expression. Early work with an engineered ZF protein directly fused to the herpes simplex viral VP16 transactivator demonstrated proof-of-principle of targeted transcriptional activation by inducing the expression of an extrachromosomal transgene in human cells (Liu et al. 1997). Subsequent studies showed that tandem repeats of VP16 were even more robust transcriptional activators than VP16 alone when linked to ZFs (Beerli et al. 1998), TALEs (Zhang et al. 2011), and nuclease-null deactivated Cas9 (dCas9) (Fig. 1E; Gilbert et al. 2013; Konermann et al. 2013; Maeder et al. 2013b; Mali et al. 2013a; Perez-Pinera et al. 2013a). Tetrameric VP16 domains (termed “VP64”) have exhibited the most widespread application as transcriptional activation domains, although larger multimers of VP16 have also been reported (Cheng et al. 2013). VP16 domains recruit cellular cofactors, such as components of the basal transcriptional machinery and chromatin remodelers (Hirai et al. 2010). Other transcriptional activation domains function similarly and have also been used in engineered transcription factors (Kim et al. 1997; Liu et al. 2001; Bikard et al. 2013; Anthony et al. 2014; Chavez et al. 2015; Konermann et al. 2015). Synergistic effects among multiple activators have been frequently observed with this class of activation domains, both when these effectors are localized in high density at adjacent sequences (Maeder et al. 2013b; Mali et al. 2013a; Perez-Pinera et al. 2013a,b) and when combined in cis or trans on multiple target sites (Cheng et al. 2013; Chakraborty et al. 2014; Gao et al. 2014; Tanenbaum et al. 2014; Konermann et al. 2015). This synergy is likely related to enhanced subunit recruitment and/or effective decreases in cofactor dissociation rates at targeted loci.

Synthetic DBPs can also function as programmable transcriptional repressors. Localization of dCas9 near transcription start sites (TSSs) can repress active gene expression (Bikard et al. 2013; Qi et al. 2013). However, when fused to repressive domains, such as the KRAB domain (Margolin et al. 1994), the inhibitory effect of ZFs (Beerli et al. 1998), TALEs (Cong et al. 2012), or dCas9 (Gilbert et al. 2013) on transcription is markedly enhanced. Transcriptional repression by these methods is often accompanied by changes in chromatin structure (Groner et al. 2010; Kears et al. 2015), which is likely a reflection of secondary KRAB-mediated recruitment of chromatin remodelers (Ying et al. 2015). Other domains have also been used for programmed transcriptional repression (Beerli et al. 1998; Snowden et al. 2002; Cong et al. 2012; Mahfouz et al. 2012; Gilbert et al. 2013; Kears et al. 2013). In contrast to artificial gene activation with effectors and recruited cofactors, repression using synthetic DBPs has not been observed to function synergistically. In addition, the degree of repression by different tools varies dramatically, even within the same DNA-targeting platform and when targeting sequences in close proximity. Additional work is needed to characterize the factors that determine the potency of gene repression, which could be related to variable targeting affinities, interactions with endogenous factors, and/or local chromatin architecture.

Next generation genome engineering: epigenome editing

The ability to readily toggle epigenetic states holds tremendous value for basic research and potentially for human therapies. Efforts aimed at editing the epigenome using synthetic DBPs are rapidly evolving (Jurkowski et al. 2015). These methods can be used to provide evidence of the causality of epigenetic marks such as DNA methylation and histone subunit modifications. Furthermore chromatin-remodeling domains fused to DBPs have also expanded our ability to modulate genomic regulatory regions. For instance, transcriptional manipulation mediated by synthetic DBPs has been most well-characterized when targeted to within 300 base pairs of TSSs. However, directed modulation of distal regulatory elements, such as enhancers, has recently been shown to be possible, albeit with varying efficacy (Gao et al. 2013, 2014; Mendenhall et al. 2013; Ji et al. 2014; Frank et al. 2015; Hilton et al. 2015; Kears et al. 2015). Enhancers represent dynamic genomic regulatory modules with differential functionality during cell lineage specification, and perturbation of enhancer function has been strongly associated with disease (Heinz et al. 2015; Roadmap Epigenomics Consortium et al. 2015). Similar to other regions of eukaryotic genomes, enhancers are controlled by
dynamic epigenetic states, including methylated DNA and post-translational modification of histone subunits (Shlyueva et al. 2014; Heinz et al. 2015). Thus, epigenome editing tools for manipulating these epigenetic modifications are critical to facilitating our understanding of the links between gene regulation, development, and disease.

High levels of S-methylcytosine (SmC) at enhancers and promoter regions are frequently correlated with transcriptional repression (Jones 2012; Schübel 2015). Initial work with designer ZFs fused to prokaryotic DNA methyltransferases demonstrated targeted methylation of DNA in vitro (Xu and Besbor 1997; McNamara et al. 2002; Nomura and Barbas 2007; Smith and Ford 2007), on extrachromosomal (Nomura and Barbas 2007) or integrated plasmid DNA (Smith and Ford 2007), and at endogenous eukaryotic targets (Carvin et al. 2003). Furthermore, targeting of mammalian DNA methyltransferases with synthetic DBPs has established that sitespecific DNA methylation at promoters can repress endogenous gene expression (Li et al. 2007; Rivenbark et al. 2012; Siddique et al. 2013). Currently, direct and targeted DNA methyltransferase activity to endogenous genes has only been applied using engineered ZF and TALE (Bernstein et al. 2015) protein scaffolds, however, it is probable that similar strategies could be adapted to CRISPR/Cas platforms.

The targeted demethylation of genomic DNA has also been used to activate gene expression using artificial DBPs. The TET family of proteins catalyzes oxidation of SmC in eukaryotic genomes, leading to reversion to unmethylated cytosine following DNA replication (Lu et al. 2015). Direct fusion of the TET1 catalytic domain to TALEs targeting regions near endogenous human genes decreased DNA methylation, leading to increased mRNA expression (Maeder et al. 2013a). In addition, the catalytic domain of murine TET2, and to a lesser extent TET1, decreased DNA methylation when targeted to human promoter regions by engineered ZF proteins (Chen et al. 2014). Artificial localization of murine thymidine DNA glycosylase, an enzyme involved in cytosine demethylation (Cortellino et al. 2011) has also been shown to decrease DNA methylation and augment gene expression from an endogenous target promoter (Gregory et al. 2013). Collectively, these results demonstrate that the targeted manipulation of DNA methylation is possible, and cytosine methylation is functionally linked to controlling gene expression.

Certain modifications on the histone subunit tails of nucleosomes are highly correlated with genomic regulatory activity (Zhou et al. 2011; Shlyueva et al. 2014; Heinz et al. 2015). In order to take advantage of this mode of transcriptional regulation and also to develop tools to better understand its roles in gene regulation, there has been a recent emphasis on targeted perturbation of histone modifications. Acetylation at lysine residues 27 and 9 of histone subunit H3 (H3K27ac and H3K9ac, respectively) are generally enriched at loci associated with high transcriptional activity such as active promoters and enhancers. A fusion of the acetyltransferase core domain of the human EP300 protein robustly activated endogenous human genes when targeted to promoter or enhancer loci using ZFs, TALEs, and dCas9 variants (Hilton et al. 2015). This activation was accompanied by enrichment for H3K27ac at a targeted promoter and at a targeted enhancer. Notably, targeted H3K27 acetylation at the well-characterized human beta-globin HS2 enhancer using dCas9 fused to the catalytic core of EP300 also led to H3K27ac enrichment and transcriptional induction from HS2-responsive promoters. Together, these results support a model in which acetylation plays a causal role in gene activation. Furthermore, this suggests that H3K27ac enrichment at human enhancers may precede and coordinate distal H3K27ac deposition. Whether this deposition occurs through physical genomic contacts and/or other endogenous factors is the subject of ongoing study. Moreover, the direct manipulation of chromatin signatures using a chromatin acetyltransferase domain appeared to be mechanistically distinct from effectors requiring other cofactors for activity, such as VP64 (Hilton et al. 2015). Thus, improvements in these programmable epigenomic modifiers may enhance the synthetic engineering of gene activation.

The acetylation of histone subunit tails can be reversed by histone deacetylases (HDACs). Recent studies have used full proteins or truncated protein domains with HDAC activity fused to ZFs (Keung et al. 2014) or TALEs (Konermann et al. 2013) to silence gene expression. Histone subunits are also dynamically regulated through methylation and demethylation of lysine residues. The catalytic regions of histone H3K9 methyltransferases EHMT2 (also known as G9A) and SUV39H1 have been found to repress transcription and alter chromatin status at targeted promoters when fused to ZFs (Snowden et al. 2002; Falahi et al. 2013; Heller et al. 2014) and TALEs (Konermann et al. 2013). Targeted H3K4 demethylation has also been applied using TALEs or dCas9 fused to the KDM1A protein (also known as LSD1) (Mendennah et al. 2013; Kearns et al. 2015), enabling the characterization of known and putative enhancers.

In addition to modulation of RNA methylation and histone residues, synthetic DBPs have also been used to manipulate chromosomal architecture. ZFs designed to artificially coordinate genomic looping between the HS2 enhancer of the globin locus control region and the beta-globin promoter activate gene expression in mouse cells (Deng et al. 2012) and similar designs can direct differential gene expression patterns between HS2 and globin genes in human and mouse cell lines (Deng et al. 2014). These results suggest that the physical interactions between enhancers and promoters can have a causal effect on gene expression. Although artificially generated chromosomal contacts have not been reported yet using TALEs or dCas9 platforms, similar approaches are likely feasible and would provide useful expansions to the genome engineering toolbox for rapid characterization of the role of chromatin conformation.

Specificity of ZFs, TALEs, and CRISPR/Cas9 systems
Understanding the target specificity of synthetic DBPs is central to their efficacy as biotechnological tools and therapeutics. Some studies using artificial ZFs fused to transcriptional effector domains indicate relatively high specificity for target gene modulation (Snowden et al. 2003). However, other results suggest widespread genomic interactions with these proteins that can lead to off-target transcriptional dysregulation (Falahi et al. 2013; Grimmer et al. 2014). Surveys of ZF nucleic specificities demonstrate that off-target effects can be prevalent (Cornu et al. 2008; Pattanayak et al. 2011), although this off-target activity may be mitigated by optimized design tools that minimize confounding factors such as context-dependent ZF domain effects (Isalan et al. 2001; Maeder et al. 2008; Sander et al. 2011; Gupta et al. 2012; Persikov et al. 2015).

As the RVDs of each synthetic TALE repeat dictate base-pair recognition, TALEs can theoretically target any genomic sequence of interest. However, engineered TALE RVDs exhibit a quantifiable variance in nucleotide recognition as well as positional effects that can lead to localization at unintended sequences (Cermak et al. 2011; Miller et al. 2011, 2015; Mali et al. 2013a; Meckler et al.
Although the frequency of generating highly active TALE nucleases is typically higher than that of ZF nucleases (Kim and Kim 2014), certain limitations exist, such as apparent difficulty targeting methylated DNA (Valton et al. 2012) and requirements for thymine bases at 5’ targeting sites (Mak et al. 2012; Lamb et al. 2013). In addition, the larger size and proclivity for recombination of repetitive sequences in TALE proteins may present difficulties in certain applications, such as viral delivery (Holkers et al. 2013), although this issue has been addressed by optimizing the codon usage of repetitive RVDs (Yang et al. 2013b). Studies of genome-wide DNA-binding, gene regulation, and chromatin remodeling suggest a high level of specificity of TALE-based transcriptional activators, although binding to off-target sites is measurable (Polstein et al. 2015).

The interaction between Cas9 and a gRNA leads to conformational changes that activate surveillance for specific target sites by Cas9 (Jinek et al. 2014; Nishimasu et al. 2014; Jiang and Doudna 2015). However, off-target interactions between dCas9 and genomic DNA have been observed in human cells, even in the absence of gRNAs (Kuscu et al. 2014; Wu et al. 2014a; O’Geen et al. 2015; Polstein et al. 2015). Although off-target binding events likely occur with dCas9-based transcriptional/epigenetic modifiers, assessments of global gene expression suggest that changes are largely restricted to the intended target sites (Gilbert et al. 2013; Perez-Pinera et al. 2013a; Hilton et al. 2015; Polstein et al. 2015). Cas9 nucleases have also been found to cause DSBs at unintended sites (Fu et al. 2013; Hsu et al. 2013; Mali et al. 2013a; Pattanayak et al. 2013), and this activity is currently thought to be related to factors including gRNA composition, chromatin accessibility, and gRNA seed/PAM sequence abundance. Several algorithms exist that allow researchers to predict potential off-target Streptococcus pyogenes gRNA binding sites and aid in optimal gRNA design (Hsu et al. 2013; Bae et al. 2014; Cradick et al. 2014; Heigwer et al. 2014; Singh et al. 2015). However, this is clearly an area where significant future research is needed.

**Applications of modern genome engineering technologies**

**Gene knockouts**

The most established application of modern genome engineering technologies is the disruption of loci through targeted nuclease activity (Fig. 2A; Urnov et al. 2010; Joung and Sander 2013; Hsu et al. 2014; Sander and Joung 2014). Genetic knockouts in eukaryotes through NHEJ-mediated disruption both in cell culture (Porheus and Baltimore 2003; Perez et al. 2008; Miller et al. 2011; Cho et al. 2013; Cong et al. 2013; Jinek et al. 2013; Mali et al. 2013b; Liao et al. 2015b) and in animals (Bibikova et al. 2002, 2003; Doyon et al. 2008; Meng et al. 2008; Cui et al. 2011; Wood et al. 2011; Hwang et al. 2013) can lead to complete loss of gene function. Moreover, genetic knockouts applied to agriculturally relevant plants and animals using genome engineering methodologies are poised to revolutionize the nutrition content and the availability of food crops and livestock (Hsu et al. 2014; Ni et al. 2014; Cryanoski 2015). In addition, these techniques to disrupt genetic information or interrogate gene function offer advantages over others, such as RNAi, which may have substantial off-target
effects and incomplete abrogation of mRNA (Shalem et al. 2015). Deletions of genomic regions are also useful for removing entire genes or portions of the genome using the concurrent action of two targeted nucleases flanking the region to be excised (Lee et al. 2010; Kim et al. 2013; Essletzbichler et al. 2014; Ousterout et al. 2015b).

The complete and precise deletion of a gene or genetic segment may obviate potential confounding factors of gene knockout with a single nuclease that might still lead to functionally active truncated or frame-shifted proteins (Shi et al. 2015). Disruption or deletion of promoter or enhancer regions could also be used to knockout or diminish gene function, and enhancer elements have been characterized and validated through such methods (Bauer et al. 2013; Li et al. 2014; Mansour et al. 2014; Zhou et al. 2014a). Deletion of boundaries between topologically associating domains can also reveal important mechanistic properties of genome structure (Nora et al. 2012; Down et al. 2014; Crane et al. 2015b; Lupíñez et al. 2015). However, deletion of genomic sequences may have disadvantageous pleiotropic effects, such as unintended alterations in the native architecture of bystander regulatory elements.

Gene knock-ins
When appropriate donor DNA repair templates are provided, HDR can lead to defined integration events such as the inclusion of epitope tags, reporter genes, and regulatory units, such as LoxP sites, at endogenous loci (Fig. 2B; Hockemeyer et al. 2009, 2011; Doyon et al. 2011a; Yang et al. 2013a). These methodologies allow for epitope-based detection of proteins and interacting partners for which high quality or specific antibodies are not available, as well as tracking of cellular proteins in real time on a single-cell basis. In addition, conditional alleles are useful in instances in which genetic knockout results in embryonic lethality or when assessments of gene function at different developmental stages and lineages are desired. HDR-mediated introduction of natural genetic variation into an isogenic background can also be used to elucidate the in vivo contributions of specific regulatory elements and DNA-interacting proteins. For instance, the targeted alteration of specific transcription factor binding site motifs in otherwise intact loci could reveal the functional contribution of transcription factor binding to regulatory element activity. This approach could also be extended to dissect cis regulatory modules in which several transcription factor binding sites are putatively involved in regulatory specificity (Hardison and Taylor 2012; Hnisz et al. 2015).

Dynamic regulation of genomic activity and conformation
In addition to nuclease-mediated disruption or deletion, genes and associated regulatory regions can also be dynamically manipulated using targeted ZF, TALE, or dCas9-based transcription factors or epigenome editing tools (Fig. 2C). This approach is particularly useful for avoiding the stochastic and cell-type- and cell-cycle-dependent DNA repair pathways involved in nuclease-mediated genome editing. Such methods can elucidate potential endogenous gene function without exogenous overexpression or permanent sequence disruption. These tools are also especially applicable in targeting the multitude of putative regulatory regions that contain epigenetic hallmarks correlated with activity in certain settings, such as differentially active enhancers (Shlyueva et al. 2014; Farh et al. 2015; Heinz et al. 2015; Roadmap Epigenomics Consortium et al. 2015; Leung et al. 2015). For example, recent efforts have identified signatures, such as DNase hypersensitivity, H3K27ac, and H3K4 methylation that are associated with active enhancers and promoters. However, validation of the potential functionality of these elements is a major scientific bottleneck to our understanding of the epigenetics of gene regulation. Selective and targeted editing or erasure of appropriate epigenetic modifications can establish the causality of respective marks in determining gene expression and may also define their relevance in organismal development and in cellular responses to distinct stimuli. Furthermore, targeted activation or suppression of regulatory loci involved in lineage specification or reprogramming could have enormous biotechnological utility (Gao et al. 2013; Chakrabarty et al. 2014; Ji et al. 2014; Chavez et al. 2015). Additionally, programmed looping to connect distal genomic regions can serve to define the function of specific genomic contacts (Deng et al. 2014). Ultimately, the most comprehensive and valuable definitions of regulatory element functionality will use several independent approaches, including sequence disruption and manipulation of activity across varied biological contexts.

Genome engineering to model disease and develop therapeutics
ZFs, TALEs, and the CRISPR/Cas9 system are also important tools for understanding and modeling disease. In Mendelian disorders, in which single gene products are implicated in disease or development, nuclease-aided disruption or deletion can be used to determine causal relationships between genes and phenotypes (Soldner et al. 2011; Toscano et al. 2013). In addition, specific SNPs or corrections can be introduced into coding regions to identify and validate variants associated with disease (Kiskinis et al. 2014; Wienert et al. 2015). SNPs in noncoding elements, such as enhancers, are also associated with many diseases; hence, similar modeling and validation methods outside of coding regions will also be extremely useful. Notably, genome engineering is unique in its ability to perturb regulatory elements in their endogenous genomic context in contrast to other tools, such as RNAi and small molecules that modulate mRNA and protein activity. Large-scale genomic models of complex diseases, based on genome-wide associations, could also be created with these strategies in order to discriminate genuine versus immaterial genetic variance (Fig. 3A; Bauer et al. 2013). Some disease phenotypes are also the result of fusion proteins generated during aberrant genomic translocations (Bunting and Nussenzweig 2013). Programmable nucleases could also be used to insert deleterious fusion proteins at endogenous loci to quantify causal effects and develop drug targets or to recapitulate chromosomal translocations to mimic disease (Maddal et al. 2014).

Programmable DBPs can also be applied for genome-wide phenotypic screening. This is especially relevant for CRISPR/Cas9-based screens, owing to the relative ease of multiplexing by simply using Cas9 in tandem with libraries of gRNAs that can be synthesized at high throughput (Shalem et al. 2015). Cas9 nuclease-based knockout screens have been recently used in combination with both positive and negative selection strategies in mammalian cell lines. These techniques have revealed genes that are essential for certain cell states and sensitivity to toxins or drugs (Koike-Yusa et al. 2014; Shalem et al. 2014; Wang et al. 2014; Zhou et al. 2014b; Chen et al. 2015). In addition, nuclease-inactivated dCas9-based repressors and activators can be used in loss-of-function and gain-of-function screens, respectively (Gilbert et al. 2014; Konermann et al. 2015). Importantly, these screening
strategies have not only validated previously implicated genes, but have also identified novel drivers of selected cellular phenotypes. The relative ease of construction of these screening platforms should make these technologies broadly useful to numerous research laboratories investigating multiple different pathologies and phenotypes.

The advent of modern genome engineering tools has also stimulated persistent and warranted optimism in the field of gene therapy. Knockout of pathological genes with Mendelian phenotypes is possible using ZFs, TALEs, and Cas9 nucleases. Similar approaches could also be used to correct deleterious copy number variation. The knockout of portions of genes or reading frame correction is also a useful strategy for pathologies in which truncated genetic variants provide amelioration, such as Duchenne muscular dystrophy (Li et al. 2015a; Ousterout et al. 2015a). Furthermore, methods for HDR-mediated gene correction and integration continue to improve, the replacement of causative SNPs at coding regions and regulatory elements may become routine (Genovese et al. 2014; Crane et al. 2015a; Hoban et al. 2015). Similarly, targeted addition of exogenous genetic payloads may also become a valuable tool to correct loss-of-function, provide dosage compensation, or create novel cellular phenotypes (Li et al. 2011; Genovese et al. 2014). Engineered DBPs may also have utility as vaccination and immunotherapeutic agents. For instance, targeted disruption of viral entry molecules, such as CCR5 for HIV, have been shown to be efficacious as a prophylactic measure (Holt et al. 2010; Tebas et al. 2014). Viral loads have also been decreased using nucleases targeted to other viral genomes (Kennedy et al. 2014, 2015; Wang and Quake 2014; Liao et al. 2015a). In addition, genome engineering efforts to enhance human T cell immunotherapy have demonstrated substantial promise (Provasi et al. 2012; Torikai et al. 2013; Beane et al. 2015) and may pave the way toward customized patient-specific prophylactics and therapies.

The therapeutic use of ZFs, TALEs, and Cas9 as artificial transcriptional/epigenetic modifiers also has tremendous potential. The precise genetic repression of detrimental genes may be a useful strategy to mitigate pathology, such as repressing the huntingtin

Figure 3. Outlook for genome engineering technologies. (A) Genome editing technologies can be used to incorporate genome-wide association study (GWAS)-identified single nucleotide polymorphisms (SNPs) into models of complex diseases and to ascertain causal disease variants. (B) Human cells can be programmed using genome engineering to produce the next generation of advanced cell therapies. (C) Techniques designed to assess potential off-target activity of programmable nucleases and transcriptional/epigenetic modifiers will facilitate technology improvements. (D) Increased control of genome engineering tools can be achieved with chemical and optogenetic regulation to develop highly articulated systems with increased specificity and dynamic properties.
gene for Huntington’s disease (Garriga-Canut et al. 2012). Similarly, targeted activation of aberrantly silenced endogenous loci may provide therapeutic benefits (Lara et al. 2012). Additionally, the manipulation of irregular epigenetic modifications or genomic contacts could be useful to prevent or correct disease states. Notably, in contrast to permanent nuclease-based sequence modification, synthetic transcriptional or epigenetic modifications are dynamic, tunable, and reversible. In certain circumstances, limited durations or transmission of artificial manipulation may be preferable. Furthermore, the versatility and multiplex capacity of these genome engineering tools combined with the recent advances in synthetic eukaryotic genetic circuit designs (Khalil et al. 2012; Sliusarzyk et al. 2012; Esvelt et al. 2013; Farzadfar et al. 2013; Daringer et al. 2014; Kabadi et al. 2014; Kiani et al. 2014; Nielsen and Voigt 2014; Nissim et al. 2014; O’Connell et al. 2014; Li et al. 2015b; Zalatan et al. 2015) may lead to the next generation of cell therapies (Fig. 3B).

Future outlook

Genome engineering tools have been widely implemented for editing and understanding eukaryotic genomes. However, technological improvements are still needed to fulfill the potential of these technologies. Ideal genome engineering tools would have completely predictable effects, lack toxicity, be easily designed and constructed, and be easily deliverable with high efficiency in vitro and in vivo. Although ZFs, TALEs, and CRISPR/Cas9 platforms have propelled the field of genome engineering, they still suffer potential limitations in all of these areas. The comprehensive characterization and optimization of targeting specificities and in vivo delivery parameters of modern genome engineering tools are arguably the most pressing concerns (Kim and Kim 2014; Cox et al. 2015).

The successful application of genome engineering in diverse organisms demonstrates that the technologies are effective with limited negative side effects in vivo, However, the translation of these tools to clinical settings requires a greater resolution of any off-target effects than has previously been possible. Some examples of the possible off-target activity of the ZF, TALE, and Cas9-based platforms have been published (Perez et al. 2008; Fu et al. 2013; Hsu et al. 2013; Mali et al. 2013a; Pattanayak et al. 2013; Guilinger et al. 2014a; Kim and Kim 2014). An important recent advance has been the development of unbiased approaches for detecting off-target effects with a much greater level of sensitivity than previous methods (Fig. 3C; Gabriel et al. 2011; Frock et al. 2015; Kim et al. 2015; Ran et al. 2015; Tsai et al. 2015). The use of large sets of negative control gRNAs in high-throughput CRISPR/Cas9-based screening applications is also enabling reproducible and highly quantitative analyses of off-target gRNA activity (Gilbert et al. 2014; Sanjana et al. 2014). However, despite the importance of these advances, their detection limits are still restrained by the accuracy of current DNA sequencing technologies, and this will be a critical area for improvement moving forward.

The protein engineering of nuclease domains to improve activity and specificity will also continue to be a focus area, particularly as the methods for computational design and experimental selection of proteins with novel properties improve. For example, engineered “nickases” have been developed to stimulate HDR but minimize off-target indels generated by the creation of double-strand breaks by nuclease, but the nickase-mediated HDR frequencies are generally lower (Doyon et al. 2011b; Miller et al. 2011; Kim et al. 2012; Wang et al. 2012; Mali et al. 2013a; Ran et al. 2013; Guilinger et al. 2014b; Tsai et al. 2014; Wu et al. 2014b). Protein engineering may also be used to alter the specificity of sequence recognition requirements of DNA-binding proteins, such as the directed evolution of Cas9 variants to recognize alternate PAMs (Kleinsteiver et al. 2015) or of TALEs to change the requirement of a 5′ thymine (Lamb et al. 2013). Continued efforts to generate effective guidelines for optimal gRNA design will also be extremely important for future Cas9-based technologies (Briner et al. 2014; Cho et al. 2014; Doench et al. 2014; Fu et al. 2014; Chari et al. 2015; Farboud and Meyer 2015; Singh et al. 2015; Xu et al. 2015). For example, the development of algorithms that incorporate biological phenomena into optimal gRNA design are an important area for future work (Singh et al. 2015). Finally, the continued development of targeted recombinases and transposases could eventually supplant nuclease-based genome editing if the challenges of efficiently and specifically targeting highly active versions of these proteins to new sites can be overcome.

The specificity of the genome engineering technologies is also pertinent for transcriptional and epigenetic modification using these tools. Although genome-wide measurements of gene expression and chromatin structure suggest low frequencies of off-target effects (Hilton et al. 2015; Polstein et al. 2015), a thorough and controlled assessment of unintended epigenetic modifications is still necessary for many of the newly described effector domains. The duration and heritability of these synthetically deposited modifications should also be the subject of future investigation (Kungulovski et al. 2015). Optimal concentrations and spatiotemporal control by chemical or optogenetic regulation will also aid in mitigating any inadvertent effects (Fig. 3D); Polstein and Gersbach 2012, 2015; Konermann et al. 2013; Davis et al. 2015; Nihongaki et al. 2015; Wright et al. 2015; Zetsche et al. 2015). In addition, the direct delivery of purified ZFs, TALEs, or preassembled Cas9/gRNA complexes can minimize off-target activity (Gaj et al. 2012; Chen et al. 2013b; Ru et al. 2013; Kim et al. 2014; Ramakrishna et al. 2014; Liu et al. 2015; Zuris et al. 2015), which may further safeguard against aberrant effects.

Despite these areas of potential optimization and improvement, genome engineering technologies are clearly ready to have a significant impact on genomics and medicine. Functional characterization of the epigenetic modifications associated with gene regulation is now possible using these tools. This will allow high resolution functional annotation and indexed categorization of genomic regulatory elements. The assignment of functional data to these regulatory regions will also benefit from analyses across lineages and cell states, thereby providing regulatory atlases across genomic space and organismal development. Ongoing improvements in delivery systems for primary cells and tissues will continue to facilitate this work (Kabadi et al. 2014; Ran et al. 2015; Zuris et al. 2015). Similarly, the manipulation of epigenetic states at specific loci will also allow determinations of causal effects of these marks and the role of epigenomic signatures at specific regulatory elements in disease and lineage specification.

The precise perturbation of genomic contact points may also be useful to determine the functionality of physical connections and to understand the reported stochasticity associated with these interactions (Kind et al. 2013; Nagano et al. 2013). Imaging of genomic loci using engineered DBPs is also possible (Chen et al. 2013a; Miyanari et al. 2013), and this could provide an independent approach to demonstrate colocalization of genomic regions in real time. In addition, directed localization of specified genomic regulatory elements to nuclear regions associated with repression,
such as the nuclear lamina (Amendola and van Steensel 2014; Pombo and Dillon 2015), could aid in validating the functional relation of nuclear subcompartamentalization. Therefore the challenge of mapping the complex and dynamic four-dimensional genome is positioned to be a major future research area enabled by these new genome engineering technologies.

The functional characterization of the expanding sets of putative regulatory regions is nontrivial but will be facilitated by modern genome engineering. Likewise, determination of the causal variants in complex non-Mendelian diseases could be accomplished using precise recapitulation of the variants in otherwise isogenic cell cultures and animal models. Developing custom genomic models will continue to enable drug development and models of potential drug resistance (Kasap et al. 2014; Smurny et al. 2014). This will facilitate rapid identification of drugs with therapeutic efficacy and realize the potential of personalized and precision medicine by connecting genomics, therapeutic targets, and disease phenotypes. Modern genome engineering platforms are now established as indispensable research tools for diverse areas of biotechnology, and promising areas for their direct and indirect application to improving human health are rapidly expanding. This new era of genomic understanding is likely to continue to create new possibilities for genome research for the foreseeable future.

Competing interest statement
I.B.H. and C.A.G. are inventors on patent applications related to genome engineering. C.A.G. is a scientific advisor to Editas Medicine, a company engaged in therapeutic development of genome engineering technologies.

Acknowledgments
We thank Gregory Crawford and Timothy Reddy for helpful discussions on the topics of this Perspective and Matthew Gemberling and Sandi Wong for critical reading of the manuscript and helpful comments. This work was supported by US National Institutes of Health (NIH) grants R01DA036865, U01HG007900, R21AR065956, P30AR066527, U54HG050050, and an NIH Director’s New Innovator Award (DP2OD008586), and a National Science Foundation (NSF) Faculty Early Career Development (CA-REEr) Award (CBET-1151035).

References
Akopian A, He J, Boocock MR, Stark WM. 2003. Chimeric recombinases with designed DNA sequence recognition. Proc Natl Acad Sci USA 100: 8688–8691.
Amendola M, van Steensel B. 2014. Mechanisms and dynamics of nuclear lamina-genome interactions. Curr Opin Cell Biol 26: 61–68.
Anders C, Niewoehner O, Duerst A, Jinek M. 2014. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature 513: 569–573.
Anthony K, Morse A, Zhang X. 2014. Activation of silenced cytokine gene promoters by the synergistic effect of TBP-TALE and VP64-TALE activators. PLoS One 9: e95790.
Bae S, Park J, Kim JS. 2014. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 30: 1473–1475.
Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero D, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315: 1709–1712.
Bauer DE, Kamran SC, Lessard S, Xu J, Fujisawa Y, Lin C, Shao Z, Canver MC, Smith EC, Pinello L, et al. 2013. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science 342: 253–257.
Beane JD, Lee G, Zheng Z, Mendel M, Abarde-Daga D, Bharathian M, Black M, Gandhi N, Yu Z, Chandran S, et al. 2015. Clinical scale zinc finger nuclease-mediated gene editing of PD-1 in tumor infiltrating lymphocytes for the treatment of metastatic melanoma. Mol Ther 23: 1380–1390.
Beerli RR, Segal DJ, Driever W, Barbash CI. 1998. Chromatin gene expression at will: specific regulation of the mtkb-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc Natl Acad Sci USA 95: 14662–14663.
Bhakta MS, Henry IM, Ousterout DG, Das KT, Lockwood SH, Meckler JF, Wallen MC, Zykovich A, Yu Y, Leo H, et al. 2013. Highly active zinc-finger nucleases by extended modular assembly. Genome Res 23: 530–538.
Bibikova M, Golic M, Golic KG, Carroll D. 2002. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161: 1169–1175.
Bibikova M, Beumer K, Trautman JK, Carroll D. 2003. Enhancing gene targeting with designed zinc finger nucleases. Science 300: 746.
Bikard D, Jiang W, Samai P, Hochsand A, Zhang F, Marraffini LA. 2013. Programmable repression and activation of bacterial gene expression using engineered CRISPR-Cas system. Nucleic Acids Res 41: 7429–7437.
Bitinjane I, Wah DA, Aggarwal AK, Schildkraut I. 1998. FokI dimerization is required for DNA cleavage. Proc Natl Acad Sci USA 95: 10570–10575.
Braun P, Dresser DM, Shochet SM, Landgraf A, Hahn S, Kay S, Lahaye T, Nicke A, Schier AS, Bonus U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326: 1509–1512.
Boissel S, Jarjour J, Astrapkina A, Adey A, Goubele D, Chudapeau P, Shendure J, Stoddard BL, Certo MT, Baker D, et al. 2014. MegaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. Nucleic Acids Res 42: 2591–2601.
Britten R, Cashmore A, Segré K, Storlach E, Nye CH, Haurwitz BE, Serfling DJ, Dreier B, Barbash CI. 1998. Towards controlling gene expression with designed zinc finger nucleases. Science 280: 443–445.
Carlson DF, Tan W, Lilloo SG, Stveraková D, Proudfoot C, Christian M, Voytas DF, Long CK, Whitelaw CB, Fahrenkrug SC. 2012. Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci USA 109: 17382–17387.
Carroll D. 2014. Genome engineering with targetable nucleases. Annu Rev Biochem 83: 409–439.
Carvin PD,uard JD, Kladde MP. 2003. Site-selective in vivo targeting of cytosine-S DNA methylation by zinc-finger proteins. Nucleic Acids Res 31: 6493–6501.
Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Soma NV, Bogdanov AJ, Voytas DF. 2011. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39: 682.
Chakraborty S, Ji H, Kabadi AM, Gersbach CA, Christoforou N, Leong KW. 2014. A CRISPR/Cas9-based system for reprogramming cell lineage specification. Stem Cell Rep 3: 940–947.
Chapman JR, Taylor MR, Boulton SJ. 2012. Playing the end game: DNA douplication in vivo. Science 335: 1060:–1068.
Chen Z, Jaafar L, Agyekum DG, Xiao H, Wade MF, Kumaran RI, Spector DL, Smith EC, Pinello L, et al. 2013. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science 342: 253–257.

Chen H, Kazemier HG, de Groote ML, Rutters MH, Xu GL, Rots MG. 2014. Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human iCM-i5 promoter. *Nucleic Acids Res* **42**:1563–1574.

Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, Scott DA, Song J, Pan JQ, Weissman JS et al. 2015. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* **160**:1246–1260.

Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, Rangarajan S, Shivailla CS, Dodan DB, Jaenisch R. 2013. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res* **23**:1163–1171.

Cho SW, Kim S, Kim JM, Kim JS. 2014. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease and nickases. *Genome Res* **24**:132–141.

Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF. 2010. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**:757–761.

Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, Kühn R. 2015. Increasing the efficiency of homology-directed repair for CRISPR-Cas9 interventions: precise gene editing in mammalian cells. *Nat Biotechnol* **33**:543–548.

Cong L, Zhou R, Kuo YC, Cunniff L, Zhang F. 2012. Comprehensive investigation of natural TALE DNA-binding modules and transcriptional repressor domain architectural effects on CRISPR-Cas systems. *Science* **339**:819–823.

Dekker J, Marti-Renom MA, Mirny LA. 2013. Exploring the three-dimensional organization of the genome in centromeric heterochromatin in mouse ES cells. *Cell Stem Cell* **4**:569–577.

Derech EA, Chylinski K, Shuman CM, Gonzales K, Ciofani A, Iishita M, Jiang L, Wei M, Saprza D, et al. 2011. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-nickases. *Nature* **473**:339–343.

Doyen JM, Fan ZP, Hnisz D, Ben G, Abraham BJ, Zhang LN, Weintzaub AS, Schuler J, Lee TI, Zhao K, et al. 2014. Control of cell identity genes occurring in insulated neighborhoods in mammalian chromosomes. *Cell* **159**:374–387.

Doyon JB, Zeitzler B, Cheng J, Cheng AT, Cheroni JM, Santiato Y, Lee AH, Vo TD, Doyon Y, Miller JC, et al. 2011a. Rapid and efficient clathrin-mediated endocytosis revealed in genome-edited mammalian cells. *Nat Cell Biol* **13**:313–337.

Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Umov FD, Gregory PD, Holmes MC. 2011b. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods* **8**:74–79.

Doudna JA, Charpentier E. 2014. Genome editing. The new frontier of genome engineering. *Nature* **489**:53–58.

Elszécsbíchter P, Konopka T, Santoro F, Chen D, Gapp BV, Kralovics R, Brummelkamp TR, Nijman SM, Schuerch FC. 2014. Mepabase-scale depletion using CRISPR/Cas9 to generate a fully haploid human cell line. *Genome Res* **24**:2059–2065.

Estvelt KM, Mali P, Brabf JL, Moosburner M, Young SJ, Church GM. 2013. Ortholog CRISPR-Cas9 proteins for RNA-gated gene regulation and editing. *Nat Methods* **10**:1116–1121.

Filet C, Huisman C, Kazemier HG, van der Vlies P, Kok K, Hosapers GA, Rots MG. 2013. Towards sustained silencing of HER2/neu in cancer by epigenetic targeting of its promoter. *Cancer Res* **73**:1029–1039.

Frick RL, Hu J, Meyers BM. 2015. Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. *Genetics* **199**:959–971.

Farboud B, Meyer BJ. 2015. Dynamic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. *Genetics* **199**:959–971.

Farboud B, Marson A, Zhu J, Kleiniewietfeld M, Housley WJ, Beik S, Shores N, Whitton H, Ryan RJ, Shishkin AA, et al. 2015. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**:337–343.

Frock RL, Perli SD, Lu TK. 2013. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *Acad Soc Biol* **2**:604–613.

Frank CL, Liu F, Wijayatunge R, Song L, Biegler MT, Yang MG, Vockley CM, Sati A, Gersbach CA, Crawford GE, et al. 2015. Regulation of chromatin accessibility and Zic binding at enhancers in the developing cerebellum. *Nat Neurosci* **18**:647–656.

Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. 2015. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol* **33**:179–186.

Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Sander JD. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**:822–826.

Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. 2014. Improving CRISPR-Cas9 nuclease specificity by truncated guide RNAs. *Nat Biotechnol* **32**:279–284.

Gabriel R, Lombardo A, Arens A, Müller JC, Genovese P, Schiroli G, Escobar G, Di Tomaso T, Firrito C, Calabria A, Moi D, Doyen JM, Fan ZP, Hnisz D, Ben G, Abraham BJ, Zhang LN, Weintzaub AS, Schuler J, Lee TI, Zhao K, et al. 2014. Control of cell identity genes occurring in insulated neighborhoods in mammalian chromosomes. *Cell* **159**:374–387.

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, et al. 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol* **26**:702–708.

Doyon JB, Zeitzler B, Cheng J, Cheng AT, Cheroni JM, Santiato Y, Lee AH, Vo TD, Doyon Y, Miller JC, et al. 2011a. Rapid and efficient clathrin-mediated endocytosis revealed in genome-edited mammalian cells. *Nat Cell Biol* **13**:313–337.

Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Umov FD, Gregory PD, Holmes MC. 2011b. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods* **8**:74–79.

The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**:57–74.

Cathomen T. 2008. DNA-binding specificity is a major determinant of adenovirus type 5 major late gene in cystic fibrosis induced pluripotent stem cells. *Stem Cell Rep* **678

Dekker J, Marti-Renom MA, Mirny LA. 2013. Exploring the three-dimensional organization of the genome in centromeric heterochromatin in mouse ES cells. *Cell Stem Cell* **4**:569–577.

Derech EA, Chylinski K, Shuman CM, Gonzales K, Ciofani A, Iishita M, Jiang L, Wei M, Saprza D, et al. 2011. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-nickases. *Nature* **473**:339–343.

Doyen JM, Fan ZP, Hnisz D, Ben G, Abraham BJ, Zhang LN, Weintzaub AS, Schuler J, Lee TI, Zhao K, et al. 2014. Control of cell identity genes occurring in insulated neighborhoods in mammalian chromosomes. *Cell* **159**:374–387.
Functional genomics with genome engineering

Kiskinis E, Sandoe J, Williams LA, Boultling G, Moccia R, Wainger BJ, Han S, Peng T, Thams S, Mikkilineni S, et al. 2014. Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. Cell Stem Cell 14: 781–795.

Kleinsteuber NP, Jin H, Tsiqou, Toporov NG, Nguyen NT, Zhong Z, Gonzales AP, Li Z, Peterson RT, Yeh J, et al. 2015. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature 523: 481–485.

Koike-Yusa H, Li Y, Tan EF, Velasco-Herrera Mdel C, Yusa K. 2014. Genome-wide repressive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol 32: 267–273.

Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, et al. 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9. Nat Biotechnol 33: 1123–1134.

Liu J, Gaj T, Wallen MC, Barbas CF III. 2015. Improved cell-penetrating zinc-finger nuclease proteins for precision genome engineering. Mol Ther Nucleic Acids 4: e232.

Lu X, Zhao BS, He C. 2015. TET family proteins: oxidation activity, interacting molecules, and functions in diseases. Chem Rev 115: 2225–2239.

Lupuñez DG, Kraft K, Heinrich V, Krawitz P, Brackt S, Kobert K, Horn D, Kayserili H, Opitz JM, Laxova R, et al. 2015. Disruptions of topological chromatin domains cause pathogenic reorientation of gene-expression interactors. Cell 161: 1012–1025.

Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han YC, Ogrodowicz P, Crippa A, Rekhtman N, de Stanchina E, et al. 2014. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR-Cas9 system. Nat Methods 11: 423–427.

Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, et al. 2008. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modulation. Mol Cell 31: 294–301.

Maeder ML, Angstman JF, Richardson ME, Linder S, Casio VM, Tsai SQ, Ho QH, Sander JD, Reynor D, Bernstein BE, et al. 2013a. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat Biotechnol 31: 1137–1142.

Maeder ML, Linder S, Casio VM, Fu Y, Ho QH, Joung JK. 2013b. CRISPR RNA-guided activation of endogenous human genes. Nat Methods 10: 977–979.

Maeder ML, Linder S, Reynor D, Angstman JF, Fu Y, Sander JD, Joung JK. 2013c. Robust, synergetic regulation of human gene expression using TALE activators. Nat Methods 10: 243–245.

Mahfouz MM, Li L, Plieck M, Fang X, Mansour H, Bangarusamy DK, Zhu JK. 2012. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. Plant Mol Biol 78: 311–321.

Mak AN, Bradley P, Cermads RA, Bogdanove AJ, Stoddard BL. 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 335: 716–719.

Mangahas KS, Haft DH, Barraganci R, Brouns SJ, Charpentier E, Horvath P, Meineau S, Mojica FJ, Wolf YI, Yakunin AF, et al. 2011. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 9: 467–477.

Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM. 2013a. Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31: 833–838.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013b. RNA-guided human genome engineering via Cas9. Science 338: 823–826.

Mansour MR, Abraham BJ, Anders L, Berezovskaya A, Gutierrez A, Durbin AD, Etchin J, Lawton L, Sallan SE, Silverman LB, et al. 2014. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science 346: 1373–1377.

Margolin AF, Friedman JD, Meyer WK, Vissing H, Thiesen HJ, Rauscher FJ III. 1994. Krüppel-associated boxes are potent transcriptional repression domains. Proc Natl Acad Sci USA 91: 4509–4513.

Maruyama M, Tsukamoto M, Michigan MC, Bilate AM, Ingram JR, Ploegh HL. 2013. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol 33: 538–542.

McNamara AR, Hurd PJ, Smith AE, Ford KG. 2002. Characterisation of site-specific zinc-finger nuclease architectures for highly specific genome editing. Nat Biotechnol 20: 81–89.

Nucifora GJ, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, Zhu JK. 2012. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. Plant Mol Biol 78: 311–321.

Ogrodowski P, Crippa A, Rekhtman N, de Stanchina E, et al. 2014. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. Plant Mol Biol 78: 311–321.

Peng T, Thams S, Mikkilineni S, et al. 2014. Pathways disrupted in human embryonic stem cells using zinc-finger nuclease. Genome Res 20: 179–188.

Poh CF, Xie Z, Wang K, Xia Z, Guo Z, et al. 2015. A novel mechanism of cell death due to zinc-finger nuclease mediated DNA double strand breaks. Cell Death Dis 6: e1613.

Poh CF, Xie Z, Wang K, Xia Z, Guo Z, et al. 2015. A novel mechanism of cell death due to zinc-finger nuclease mediated DNA double strand breaks. Cell Death Dis 6: e1613.
Miyano Y, Ziegler-Bitling C, Torres-Padilla ME. 2013. Live visualization of chromatin dynamics with fluorescent TALEs. Nat Struct Mol Biol 20: 1321–1324.

Mojica FJ, Diez-Villaseñor C, Garcia-Martínez J, Almendros C. 2009. Short mobile sequences determine the targeting specificity of the prokaryotic CRISPR defense system. Nature 458(7239): 733–740.

Morbitzer R, Romer P, Boch J, Lahaye T. 2010. Regulation of selected genome loci using de novo-engineered transcription activator-like effector transcription factors. Nat Methods 7(12): 1011. 1578:1671–1672.

Moscou MJ, Bogdanove AJ. 2009. A simple cipher governs DNA recognition by TAL effectors. Science. 326: 1501.

Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A, Fraser P. 2013. Single-cell Hi-C reveals cell-to-cell variability in the human genome. Nature 502: 59–64.

Ni W, Qiao J, Hu S, Zhao X, Regouski M, Yang M, Polejaeva IA, Chen C. 2014. Efficient gene knockout in goats using CRISPR/Cas9 system. PLoS One 9: e106718.

Nielsen AA, Voigt CA. 2014. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. Mol Syst Biol 10: 763.

Nishimura Y, Yamamoto S, Kawanoue M, Iwai Y, Ito T, Ishitani R, Zhang F, Nureki O. 2014. Crystal structure of Cas9 in complex with guide DNA and target RNA. Cell 155: 935–949.

Nisim L, Peril SH, Eldor A, Perona P, Li TK. 2014. Programmable and recoverable regulation of gene networks with an integrated RNA and CRISPR/Cas9 toolkit in human cells. Mol Cell 54: 698–710.

Norum W, Barbas CF III. 2007. In vivo site-specific DNA methylation with a designed sequence-enabled DNA methylase. J Am Chem Soc 129: 8676–8677.

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piñol T, van Berkum NL, Meisig J, Sedat J, et al. 2012. Spatial partitioning of the genome landscape of the X-chromosome inactivation centre. Nature 485: 381–385.

O’Connell MR, Oakes BL, Sterngerb SH, East-Seletsy A, Kaplan M, Doudna JA. 2014. Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature 516: 263–266.

O’Gara H, Hennig WH, Blaksta MS, Meckler JF, Segal DJ. 2015. A genome-wide analysis of Cas9 binding specificity using ChiP-seq and targeted sequence capture. Nucleic Acids Res 43: 3389–3404.

Ousterout DG, Kocak DD, Adler AF, Kabadi AM, Polstein LR, Amagasa T, Horn MJ, et al. 2013. Correction of dystrophin expression in human cells by using a designed exon 51 skipping Cas9 fusion protein with increased DNA cleavage activity. Mol Ther 21: 523–532.

Pattanayak Y, Ramirez CI, Joung JK, Liu DR. 2011. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods 8: 765–770.

Pattanayak Y, Lin S, Guillinger JP, Ma E, Doudna JA, Liu DR. 2013. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed specificity. Nature 494: 59–64.

Pavlitch NP, Pabo CO. 1991. Zinc-finger–DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. Science 252: 809–817.

Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A, Fraser P. 2012. Engineered T cell specificity to treat cancer-induced leukemias by zinc finger nucleases and lentiviral gene transfer. Nat Med 18: 807–815.

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152: 1173–1183.

Ramakrishna S, Suri NK, Bento JL, Gopalappa R, Lee SK, Kim H. 2014. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Nucleic Acids Res 42: 1020–1027.

Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, et al. 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 158: 1380–1389.

Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kitz AJ, Zetsche B, Zhang F, Wu X, Gerstle JS, Ouyang Z, et al. 2013. In vivo genome editing using Staphylococcus aureus Cas9. Nature 508: 186–191.

Rivenbark AG, Stolzenburg SL, Beltran AS, Yuan X, Rots MG, Strahl BD, Blancfort P. 2012. Epigenetic reprogramming of cancer cells via target-specific DNA methylation. PLoS One 7: e350–360.

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 520: 317–330.

Römer P, Hahn S, Jordan T, Strauss B, Bonas U, Lahaye T. 2007. Plant pathogen recognition mediated by promoter activation of the pea B3 resistance gene. Science 318: 645–648.

Rouset P, Smith F, Jasin M. 1998. Induction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 14: 8096–8106.

Ru R, Yao Y, Yu S, Yin B, Xu W, Zhao S, Qin L, Chen X. 2013. Targeted gene silencing in human induced pluripotent stem cells by penetrating TALENs. Cell Regen (Lond) 2: 5.

Sander JD, Joung JK. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32: 347–355.

Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtis SJ, Blackburn J, Thuloub-Kabeagy S, Qi Y, et al. 2011. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8: 67–69.

Sanjana NE, Shalem O, Zhang F. 2014. Improved vectors and genome-wide libraries for CRISPR/Mak screening. Nat Methods 11: 783–784.

Schubeler D. 2015. Function and information content of DNA methylation. Nature 517: 321–326.

Shalem O, Sanjana NE, Hartienan S, Hsu P, Scott DA, Mikkelsen TS, Heckl D, Bilenky M, Yen A, Heravi-Moussavi A, et al. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343: 84–87.

Shalem O, Sanjana NE, Zhang F. 2015. High-throughput functional genomics using CRISPR-Cas9. Nat Rev Genet 16: 299–311.

Shi J, Wang E, Milazzo JP, Wang Z, Wang Z, et al. 2015. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat Biotechnol 33: 661–667.

Shilyueva D, Stampel G, Stark A. 2014. Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet 15: 272–296.

Siddiqua AN, Nunnia S, Rajavele A, Zhang Y, Jurkowska RZ, Reinhardt R, Rots MG, Ragozin S, Jurkowska TP, Jeltsch A. 2013. Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3l single-chain fusion protein with increased DNA methylation activity. J Mol Biol 425: 479–491.

Singh RR, Kuscu C, Quinlan A, Qi Y, Adli M. 2015. Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. Nucleic Acids Res doi: 10.1093/nar/gkv269.

Slusarczyk AL, Lin A, Weiss R. 2012. Foundations for the design and implementation of synthetic genetic circuits. Nat Rev Genet 13: 406–420.

Smith AE, Ford KG. 2007. Specific targeting of cytosine methylation to DNA sequences in vivo. Nucleic Acids Res 35: 740–754.

Smithies O, Gregg R, Boggs SS, Koralewski MA, Kucherlapati RS. 1985. Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. Nature 317: 230–234.

Smurnyy Y, Cai M, Wu H, McWhinnie E, Taliarico JA, Yang Y, Feng Y. 2014. DNA sequencing and CRISPR-Cas9 gene editing for target validation in mammalian cells. Nat Chem Biol 10: 623–625.

Snowden AW, Gregory PD, Case CC, Pabo CO. 2002. Gene-specific targeting of Hox9 methylation is sufficient for initiating repression in vivo. J Biol Chem 277: 2159–2166.
Snowden AW, Zhang L, Urnov F, Dent C, Jounvenot Y, Zhong X, Rebar Ej, Jamieson AC, Zhang HS, Tan S, et al. 2003. Repression of vascular endothelial growth factor A in glioblastoma cells using engineered zinc finger transcription factors. Cancer Res 63: 8968–8976.

Sternberg SH, Redding SJ, Kreek M, Greene EC, Dou J, JA. 2014. DNA integration by the CRISPR RNA-guided endonuclease Cas9. Nature 507: 62–67.

Stoddard BL. 2014. Homing endonucleases from mobile group I introns: discovery to genome engineering. Mol DNA 8: 7.

Szczepek M, Brondani V, Büchel J, Serrano L, Segal DJ, Cathomen T. 2007. A protein-based transposon insertion by engineered ZF- DNA-binding domains. Mol Ther 15: 1852–1861.

Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. 2014. A protein–DNA complex engineered to generate universal cells from allogeneic donors. Cell 159: 635–646.

Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, Soprati SK, Surisky RT, Giedlin MA, Nicholl G, et al. 2014. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 370: 901–910.

Thomas KR, Folger KR, Capecchi MR. 1986. High frequency targeting of murine transcription. Nature 321: 1316–1322.

Wang J, Friedman G, Doyon Y, Wang NS, Li CJ, Miller JC, Hua KL, Yan J, Babiarz JE, Gregory PD, et al. 2012. Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. Genome Res 22: 1316–1326.

Wang T, Wei JJ, Sabatini DM, Lander ES. 2014. Genetic screens in human cells using the CRISPR-Cas9 system. Science 343: 80–84.

Xu GL, Bestor TH. 1997. Cytosine methylation targeted to pre-determined sequences. Nat Genet 17: 376–378.

Ying Y, Yang X, Zhao K, Mao J, Kuang Y, Wang Z, Sun R, Fei J. 2015. The development of CRISPR–Cas9: a decade on. Nat Rev Genet 16: 683–697.

Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, Tsai YY, Huang Y, Akache B, Kay MA. 2007. Site-directed transposon insertion by engineered ZFN and TALEN. Science 318: 1499–1503.

Zhang J, Quake SR. 2014. RNA-guided endonuclease provides a therapeutic strategy to cure latent herpesvirus infection. Proc Natl Acad Sci USA 111: 13157–13162.

Zhou HY, Katsman Y, Dhaliwal NK, Davidson S, Macpherson NN, Sakthidevi M, Collura F, Mitchell JA. 2013. A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. Genes Dev 28: 2699–2711.

Zuris JA, Thompson DB, Shi Y, Guilinger JP, Bessen JL, Hu JH, Maeder ML, Joang JK, Chen ZY, Liu DK. 2015. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotechnol 33: 73–80.