

Dissecting Neural Function Using Targeted Genome Engineering Technologies

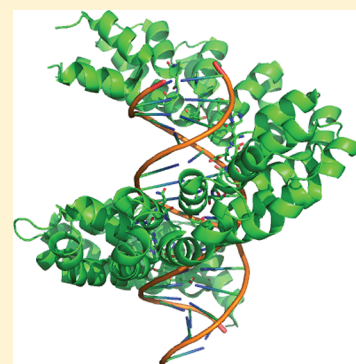
Patrick D. Hsu^{†,‡,§} and Feng Zhang^{*,‡,§}

[†]Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, United States

[‡]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States

[§]McGovern Institute for Brain Research, MIT Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

ABSTRACT: Designer DNA-binding proteins based on transcriptional activator-like effectors (TALEs) and zinc finger proteins (ZFPs) are easily tailored to recognize specific DNA sequences in a modular manner. They can be engineered to generate tools for targeted genome perturbation. Here, we review recent advances in these versatile technologies with a focus on designer nucleases for highly precise, efficient, and scarless gene modification. By generating double stranded breaks and stimulating cellular DNA repair pathways, TALE and ZF nucleases have the ability to modify the endogenous genome. We also discuss current applications of designer DNA-binding proteins in synthetic biology and disease modeling, novel effector domains for genetic and epigenetic regulation, and finally perspectives on using customizable DNA-binding proteins for interrogating neural function.



KEYWORDS: TALE effector, TALEN, zinc finger, gene targeting, designer nuclease, genome engineering

Recent advances in DNA sequencing technology offer enormous potential for understanding the genetic and epigenetic basis of nervous system function as well as disease processes. However, the detailed mechanisms of cognitive function and neuropsychiatric diseases remain poorly understood. In order to study cellular programs that orchestrate brain processes, tools that facilitate precise editing or regulation of the genome would enable systematic interrogation of neural function from molecules to behavior. We discuss advances in genome engineering technologies that provide efficient insertion, removal, replacement, or regulation of genomic sequence in a targeted manner.

By tethering different effector domains to engineered DNA-binding proteins of programmable specificity, such as TALE effectors (TALEs) or zinc finger proteins (ZFPs), it is now possible to achieve highly specific gene targeting or modulation.^{1–3} Designer nucleases, which couple the catalytic domain of restriction endonucleases to a DNA binding domain,^{4,5} have the potential to introduce specific mutations or desired transgenes almost anywhere within the genome. Targeted nucleases have been used in several gene therapy studies in mice, successfully treating diseases such as hemophilia or HIV in cellular and animal models.^{6,7} Such experiments suggest the possibility for similar strategies in the brain, from monogenic neurological disease to disorders with more complicated genetic contributions. Other functional variants of TALEs and ZFPs include transcriptional activators and repressors,^{2,8,9} opening the door to probe the epigenetic dynamics underlying brain function. Here, we review developments in targeted nuclease engineering with a focus on the

emerging TALE technology, followed by a brief examination of current applications. Finally, we will discuss perspectives on applications of designer binding proteins for understanding neurological processes and disease mechanisms.

■ NEED FOR PRECISION GENOME ENGINEERING TOOLS

A variety of genome manipulation techniques have been developed over the past decades to reverse engineer the genetic contributions to cellular phenotype, including retrovirus-mediated transgene integration, site-specific recombinases, and gene targeting through homologous recombination (HR). Retrovirus-mediated transgene integration is carried out through random or semirandom insertion into hotspots within the genome, and can cause insertional mutagenesis that may confound experimental conclusions.¹⁰ Site-specific recombinases rely on the introduction of exogenous recombinase recognition sites into the targeted locus, which is undesirable.¹¹ Finally, traditional gene targeting relies on HR, which suffers from a low absolute efficiency. Successful transgene insertion typically occurs in 1 out of 10⁶ to 10⁹ treated cells.¹² As a result, recombination efficiency is a bottleneck for stimulating efficient gene targeting.

Special Issue: Biological Engineering in Neuroscience

Received: July 10, 2012

Accepted: July 19, 2012

Published: July 19, 2012

Despite this, much progress has been made in mammalian systems, especially in mice, due to clever selection strategies and the availability of embryonic stem cells (ES cells). In order to guard against random integrants and yield the desired recombined product, a double positive/negative selection strategy can be used.¹³ ES-cell based gene targeting, however, is slow and expensive, and for many model organisms germ-line competent ES cell lines are not available and selection/screening strategies do not work well.

More advanced tools that can combine highly specific, efficient, and scarless gene modification would be an ideal technological advance for more effective genome engineering. The development of targeted nucleases is proving to be a versatile way to overcome deficiencies with traditional gene targeting approaches.

DESIGNER NUCLEASES

Studies employing homing meganucleases (MNs), a class of site-specific endonucleases, were able to stimulate homologous recombination by several orders of magnitude.^{14,15} MNs are able to introduce double strand breaks (DSBs) at specific target sequences and stimulate homologous recombination by activating the endogenous cellular DNA damage repair machinery (Figure 1, Table 1). However, applications involving

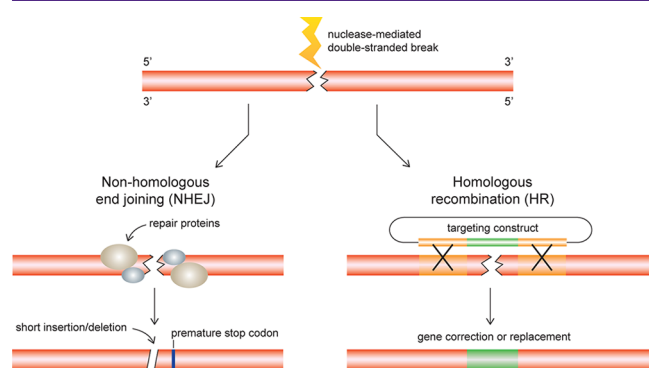


Figure 1. Designer nuclease activates endogenous DNA damage repair pathways. Designer nuclease-induced DNA double strand breaks are typically repaired by two major DNA damage repair pathways: nonhomologous end joining (NHEJ) and homologous recombination (HR). NHEJ directly ligates the ends of the broken DNA in an error-prone process that can introduce a short insertion or deletion at the repair site, leading to missense or nonsense mutations and loss of gene function if targeted to protein coding regions. HR allows for finer manipulation by introducing a targeting construct flanked by long homology arms in order to direct recombination to the DSB site.

MNs for targeted genome editing have remained limited due to the low probability of finding MN-binding sites at the target genomic locus.¹⁶ Altering the substrate specificity of MNs has proved challenging, as their DNA binding and nuclease catalytic

domains are coupled, making it difficult to reengineer specificity without affecting cleavage activity.

The development of zinc finger nucleases (ZFNs) has proved to be a versatile way to address these problems with MN engineering. Separating the DNA recognition and cleavage domains of the ZFN made it possible to alter substrate specificity without impairing endonuclease cleavage. Zinc fingers (ZFs) are protein structural motifs that direct some transcription factors to their genomic targets. ZF domains can be found in tandem arrays of individual fingers, with each finger within the larger ZF protein recognizing a typically 3 bp DNA sequence in a remarkably modular manner¹⁷ (Figure 2a). As a result, it appeared that a customized array of individual zinc finger domains assembled into a ZFP could be designed to target a larger DNA sequence.

The first synthetic zinc finger nucleases were developed by fusing a ZF protein to the catalytic domain of the Type IIS restriction enzyme FokI.^{4,5} It was later realized that the FokI domain required dimerization to introduce a DSB,¹⁸ leading to the development of a paired heterodimeric nuclease design strategy that increased cleavage specificity while minimizing off-target activity (Figure 2b).¹⁹ The current generation of ZFN cleaves genome targets containing two 9–18 bp binding sites separated by a 5–6 bp spacer. This is important for nuclease specificity because cleavage should not occur at single ZFN binding sites. As a result, the overall 18–36 bp sequence specificity of a ZFN pair should be unique within even a complex eukaryotic genome.

ZFNs have been used to target many genes in a wide variety of organisms (see reviews, refs 20 and 21). However, successful ZFN design and application can often be quite difficult. The most basic design strategy, modular assembly, optimizes individual zinc fingers against target triplet DNA sequences and then links them together to target a larger sequence. This assembly method suffers from a high failure rate²² when assembled in an array, because zinc-finger domains exhibit context-dependent binding preference. Individual ZFs are affected by their neighboring modules, neighboring modules, making the selection of functional ZFNs requires an extensive and time-consuming screening process^{23,24} that also has a limited target sequence space.²⁵ Zinc fingers are also known to exhibit low levels of off-target activity, although this has been minimized through careful selection and design.^{26,27}

Similar to ZFPs, transcriptional activator-like effectors (TALEs) can also be customized to recognize specific DNA sequences. They were first discovered in plant pathogenic bacteria of the genus *Xanthomonas* and are virulence factors that are translocated into rice crops via a Type III bacterial secretion system.²⁸ Once inside the host cell, TAL effectors then upregulate host genes by binding and activating target promoters in the host genome, leading to disease symptoms such as citrus canker and bacterial blight.

Table 1. Repair Mechanisms for Double-Stranded DNA Breaks

DNA double strand breaks (DSBs) activate the intrinsic cellular DNA damage repair pathways (Figure 1) and can be repaired by either homologous recombination (HR) or non-homologous end joining (NHEJ), both of which are highly conserved mechanisms. NHEJ is an error-prone process that can introduce short insertions or deletions (indels) upon repair of the DNA lesion, which leads to mutation of the locus. If the mutation is introduced within the coding region of a gene, it can lead to a nonsense or missense mutation and knock out gene function.

HR uses a repair template homologous to the damaged site for DSB repair. In principle, one can use a recombinant donor plasmid with appropriate homology arms as a substitute template for repair, instead of the undamaged sister chromatid or homologous chromosome. HR-mediated gene editing may be more versatile than NHEJ, because it allows for more controlled gene correction. An appropriately designed exogenous template can be used to create defined insertions, deletions, or gene replacement at desired genomic loci.

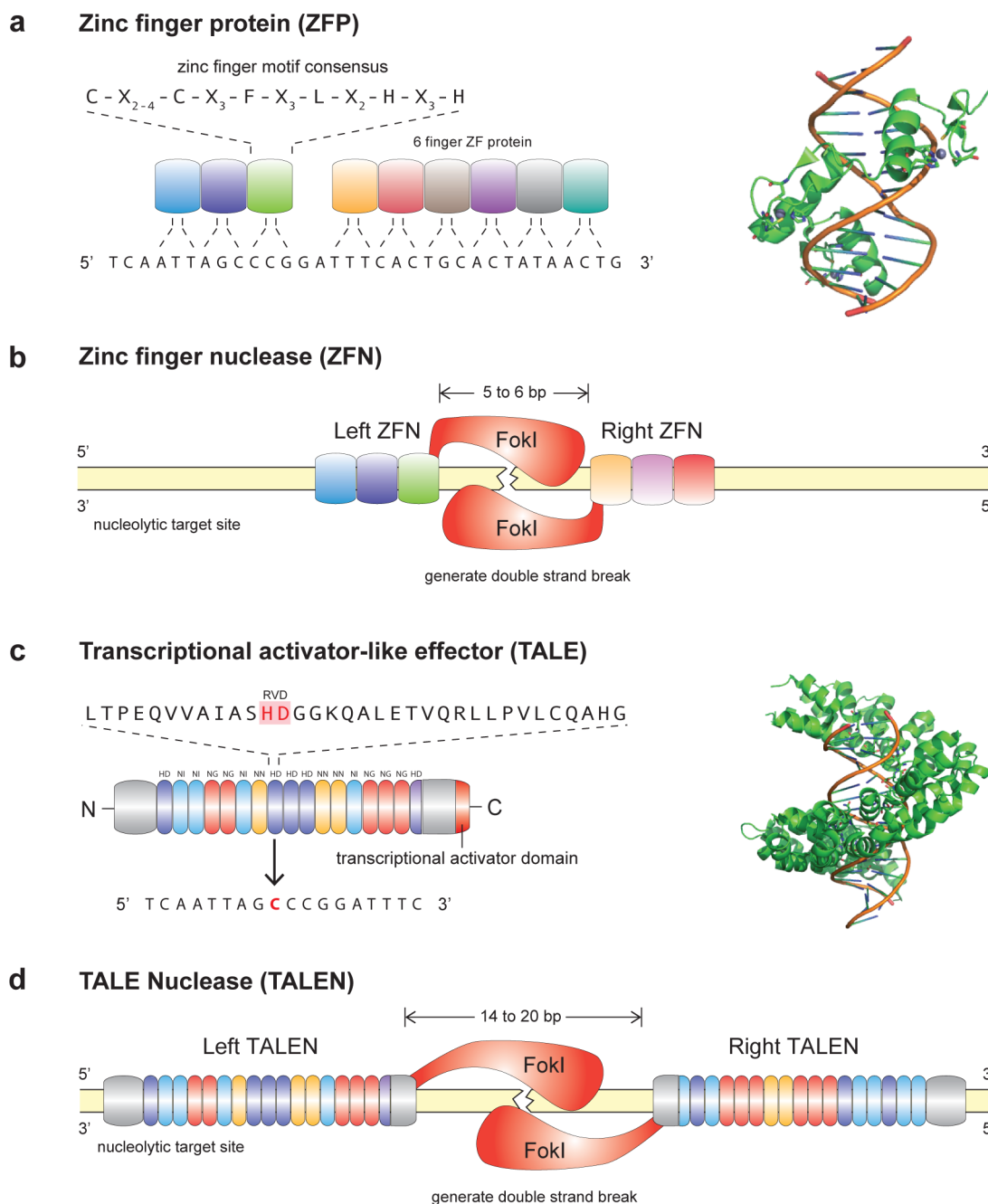


Figure 2. Designer DNA-binding domains for targeted genome manipulation. (a) A synthetic zinc finger DNA-binding domain typically consists of 3–6 zinc finger modules, with each finger within the larger ZF protein typically recognizing a 3 bp DNA sequence. Prototypical 3 and 6 finger zinc finger proteins are shown, targeting sequences of 9 and 18 bp in length, respectively. A crystal structure of three zinc finger modules in complex with DNA (right panel). (b) Zinc finger modules can be fused to the catalytic domain of the *FokI* endonuclease to generate designer zinc finger nucleases (ZFNs). A pair of ZFNs binds to their target cleavage site, allowing the *FokI* domains to dimerize and generate a double-stranded breaks (DSBs) between the pair of ZFN target sites. (c) The primary structure of a naturally occurring TAL effector in *Xanthomonas* sp., a bacterial pathogen, consists of an array of 34 amino acid tandem repeat modules flanked by nonrepetitive N- and C-terminal domains. The 12th and 13th amino acids of each repeat, known as the repeat variable diresidue (RVD), specify its DNA target base according to a simple cipher: NG = T; NI = A; HD = C; NN = G/A; NH = G. A crystal structure of TALE in complex with DNA (right panel). (d) Similar to ZFNs, the *FokI* catalytic domain can be fused to TALEs to generate designer TALE nucleases (TALENs).

Naturally occurring TALEs contain a DNA-binding domain composed of a tandem array of typically 34 amino acid modules that are largely identical (Figure 2c). Two amino acids, at positions 12 and 13 of each tandem repeat, are collectively known as a repeat variable diresidue (RVD) and have been identified to specify DNA binding.^{29,30} The target DNA sequence of a TAL effector can be deciphered based on the

sequence of RVDs and also a simple cipher correlating each RVD to its target base. Four of the most commonly occurring RVDs in natural TALEs (Asn-Ile (NI), Asn-Gly (NG), His-Asp (HD), and Asn-Asn (NN)) preferentially recognize A, T, C, and G/A respectively.³⁰ The RVD Asn-His (NH) has been recently reported to recognize guanine more specifically.⁹

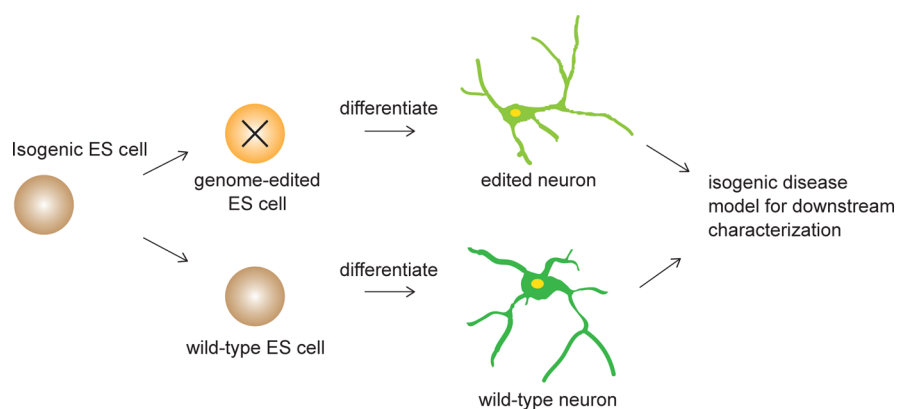


Figure 3. Designer nuclease can be used to generate isogenic cell models for probing the causal role of disease mutations. Introducing specific mutations of interest into isogenic stem cells can generate isogenic disease models, controlling for variations in genetic background. These isogenic cells can then be converted into the relevant neuron types or other cell types for downstream characterization.

By assembling specific combinations of repeat monomers with different RVDs, synthetic TALE DNA-binding domains can be constructed to target novel DNA sequences. TALEs are increasingly being adopted because their DNA binding domains are more modular and easier to customize than those of zinc finger proteins. In addition, designer TALEs can be synthesized in less than 1 week.^{2,31–35} TALEs can be fused to the FokI catalytic domain to generate TALE nucleases (TALENs) and facilitate genome editing (Figure 2d).^{3,31–33,36,37}

■ APPLICATIONS FOR DESIGNER NUCLEASES

The ability to modify the endogenous genome at will opens a wealth of opportunities for probing the genetic and epigenetic basis of cellular function and disease. Designer nucleases have already been widely adopted to generate novel animal and cellular models of diseases that more accurately model human disease genetics and mechanisms.³⁸ For instance, ZFNs and TALENs can be used to introduce the exact mutations found in patient populations into animal and cellular models to study the causal role of specific disease mutations. Additionally, targeted nucleases enable gene targeting in a significantly broader spectrum of animal species beyond the traditionally used laboratory mice, thereby giving researchers the ability to choose the most appropriate animal and cellular models for their respective studies.

Novel Transgenic Animal Models. Designer nucleases can be used to increase the efficiency of genome modification in transgenically accessible animal models as well to enable the manipulation of organisms that are difficult to manipulate using conventional gene targeting methods. For example, embryonic stem cell (ES cell) based knockout generation in rodents is notoriously time-consuming and typically requires at least 1 year of backcrossing. By introducing ZFNs or TALENs into the pronucleus of single cell mouse embryos, knockout animals can be generated in as little as 4 months.³⁹

The rat has traditionally been an intractable genetic model, which has been frustrating because it serves as an excellent model organism for complex human disease traits as well as for neurophysiological studies. Technical advances over the last 4 years have led to the capture of authentic rat ES cells^{40,41} and the subsequent production of knockout rats from modified rat ES cells.⁴² However, as in mice, ES cell targeting in rats also requires selection screening and laborious backcrossing to derive the desired genetic background. Both ZFNs and TALENs have been established as tools for robust targeted

integration and *in vivo* gene knockout in rats.^{43,44} Targeted nucleases can significantly increase the targeting efficiency and reduce the overall time required for achieving successful gene knockout.

Designer nucleases are also promising for model organisms that have not proved to be amenable to traditional methods for genetic manipulation. Customized nucleases can also be used to facilitate gene targeting in higher level mammals such as nonhuman primates, which have been an intractable model organism because monkey ES cells have so far failed to incorporate into host blastocysts.^{45,46} TALENs and ZFNs may be used to directly engineer embryos or sperm progenitor cells,⁴⁷ thus overcoming the difficulties facing ES cell-based approaches. Particularly for neuroscience and neuropsychiatric diseases studies, nonhuman primate animal models may be able to better recapitulate the disease mechanisms and behavioral sophistication manifested in human neuropsychiatric disease conditions.

Isogenic ES Cell Models. Genetic manipulation of ES and induced pluripotent stem cells (iPSCs) is a powerful technique for studying the genetic basis of diseases that affect a particular cell type, developing patient-specific therapies, and investigating developmental gene regulation. While classical gene targeting in ES cells has been limited due to low HR rates of less than 1 in 10^6 cells, the advent of targeted nucleases has led to the development of many stem cell-based applications (see ref 48 for review).

Recent advances in the genetic analysis of patient genomes have uncovered large sets of genetic mutations implicated in disease. However, the complexity of human genetic backgrounds renders the study of genetic basis of disease difficult. In addition to the locus of interest, cells derived from two distinct individuals will likely differ in over 10^6 positions. Targeted nucleases based on TALENs and ZFNs are already enabling researchers to introduce specific mutations of interest into isogenic pluripotent stem cells (PSCs) so that these isogenic cells can be converted into the relevant cell types for downstream studies (Figure 3).³⁸ By recapitulating possible deletions, duplications, or other genetic mutations that are implicated with the disease pathophysiology, targeted nucleases can be used to develop new or more phenotypically accurate models of disease.

Nuclease-mediated editing of PSCs is of particular interest because these cell lines may be differentiated into almost every cell type. To facilitate the application of PSCs in disease

modeling, TALENs and ZFNs can be used to generate cell type-specific reporter cell lines by inserting reporter genes such as fluorescent proteins into the loci of specific cell type markers. These reporter cell lines can be used in cDNA or small molecule screening assays to identify a combination of genes that are capable of directing the differentiation of stem cells into a specific cell type. Additionally, these reporter cell lines can be engineered to carry mutations implicated in particular disease states. They can then be differentiated into an appropriate cell type, while the reporter allows for tracking of cell line differentiation. This will guide the functional study of the disease mechanisms that result from specific genetic mutations in the cell type of interest.

In Vivo Gene Correction. Successful *in vivo* gene therapy has been reported with retroviral-mediated transgene insertion. Adeno-associated virus-based delivery of a functional copy of the retinoid protein RPE65 led to visual restoration in a canine model of childhood blindness,⁴⁹ while lentiviral-mediated therapy in hematopoietic stem cells halted brain demyelination in human patients.⁵⁰ However, there are important caveats with these delivery methods. All forms of viral-mediated gene replacement suffer from loss of endogenous regulation because of random or semirandom insertion of the viral vector. In some cases, this can also lead to the activation of proto-oncogenes by retrovirus enhancer activity.⁵¹

Designer nucleases targeted to endogenous loci would avoid both of these problems with viral therapy. One solution has been to pursue *ex vivo* therapeutics by performing targeted gene correction in ES and iPS cells followed by transplantation into patients. This approach to therapeutic cloning was successfully used to knockout the receptor and treat HIV in human patients,⁶ as well as to reconstitute the hematopoietic system of mouse models of sickle cell anemia.⁵² In addition to these cellular therapy methods, targeted nucleases also offer a more generalizable solution by enabling direct correction of gene mutations in affected cell *in vivo*. A recent study demonstrated ZFN-mediated insertion of the functional exons of blood coagulation factor IX into liver cells of a mouse model of hemophilia, a blood clotting disorder.⁷ This increased circulating levels of working blood clot factor, thereby restoring hemostasis.

Direct gene correction of these or similar disorders by TALENs is an exciting avenue for pursuing gene therapy in the brain. Some neurodegenerative genetic disorders are monogenic or largely caused by a small set of mutations, such as Huntington's disease, yet many of these diseases have no known cure. Gene therapy strategies may be developed to treat these neurological genetic disorders through stereotactic delivery of therapeutic vectors into specific brain regions or brain-wide gene delivery through ventricular or spinal injection. Table 2 lists several especially attractive candidates for targeted nuclease-mediated gene therapy.

Additional Applications and Considerations. Although the versatility of targeted nucleases in successful strategies for gene therapy is highly encouraging for investigating neurological function, it is not yet clear if targeted nucleases can be applied in the brain. One important consideration for gene editing within the developed brain is the DNA repair capabilities of nondividing cells such as postmitotic neurons. Given that the versatility of nuclease-mediated therapies relies on homology-directed repair, it is important to better understand the DNA break repair capabilities of mature neurons.

Table 2. Select List of Genetic Neurological Disorders

disease	classification	description
Huntington's disease (HD)	Autosomal dominant	A monogenic neurodegenerative disease caused by expansion of a CAG triple nucleotide repeat stretch in the <i>huntingtin</i> gene. HD affects the entire brain and although the disease mechanism is unclear, it is marked by degeneration of medium spiny neurons within the striatum.
Fragile-X syndrome	X-linked dominant	Similar to Huntington's disease, Fragile-X results from the expansion of a CCG trinucleotide repeat stretch within the fragile X mental retardation 1 (FMR1) gene on the X-chromosome. It is a common form of inherited mental retardation and is the leading known genetic cause of autism, contributing to an estimated 5% of cases. FMR1 is thought to regulate protein synthesis at synapses.
Angelman's syndrome	Monosomy	A particular deletion in chromosome 15 of humans leads to disruption of the ubiquitin ligase <i>Ube3a</i> gene, resulting in aberrant synaptic signaling and seizures. The paternal copy of <i>Ube3a</i> is typically silenced, and mutations in the maternal <i>Ube3a</i> gene lead to this neurological disorder.
Leukodystrophies: disorders involving myelin sheaths	Both autosomal recessive	<i>Canaan disease</i> : caused by defects in the gene <i>ASPA</i> , which codes for the enzyme aspartoacylase, which breaks down the aspartic acid derivative N-acetyl aspartate (NAA). High levels of NAA interfere with myelin sheath formation and profoundly affect neural signaling. <i>Krabbe disease</i> : caused by mutations in the <i>GALC</i> gene, which codes for galactocerebrosidase, an enzyme that metabolizes sphingolipids. Lipid buildup leads to defective myelin sheaths, degeneration of motor development, and often death.
Neurofibromatosis	Autosomal dominant	Mutations in neurofibromin 1 (NF1), a tumor suppressor gene, lead to neurofibromatosis. NF1 deficiency can result in uncontrolled cell proliferation and the formation of neurofibromas, which are nerve sheath tumors.
Tay Sachs disease	Autosomal recessive	Mutations in the <i>HEXA</i> gene affect the levels of the enzyme hexosaminidase A, which breaks down lipids. Excessive lipid buildup in the brain eventually leads to neuronal death.
Duchenne muscular dystrophy (DMD)	X-linked	This X-linked form of muscular dystrophy (MD) is caused by mutations in the <i>dystrophin</i> gene, which codes for an important structural and functional component of muscle cells. Patients with some functional copies of dystrophin tend to have Becker's MD, a milder form of Duchenne. Boys with DMD are cognitively impaired, while dystrophin-deficient mice exhibit reduced neuron number and defects in short-term memory. ⁵⁸

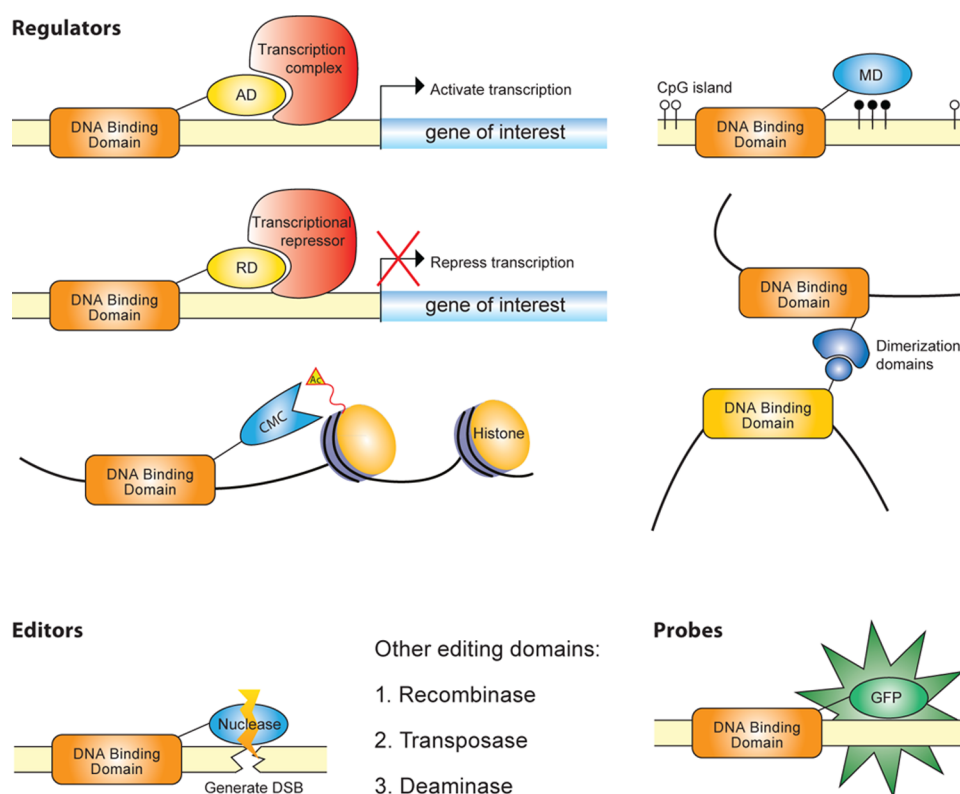


Figure 4. Novel effector domains for targeted genome perturbation. Designer DNA-binding proteins can be engineered with new effector domains to expand the existing toolkit for targeted genome perturbation. In addition to transcriptional activator and repressor domains, TALEs and ZFPs can be used to target chromatin modifying domains to specific genomic loci and facilitate histone modification, DNA methylation/demethylation, or chromosomal architecture remodeling. In addition to tools that can control genetic or epigenetic regulation, designer nucleases are being used for gene targeting. Other editing domains such as recombinases, transposases, or deaminases are alternative ways to modify genomic sequences. Finally, fluorescent proteins can be fused to DNA binding domains to illuminate chromosomal configuration. Abbreviations: transcriptional activation domain, AD; transcriptional repressor domain, RD; methyltransferase domain, MD; chromatin modifying complex, CMC.

Whereas nonhomologous end joining (NHEJ) should in principle occur in any stage of the cell cycle, homologous recombination is thought to largely occur during the S and G2 phases of the cell cycle³³ with the sister chromatid acting as the repair template. Progress of the main pathway of DSB repair is cell-cycle dependent, with end resection blockage in the G1 stage. End resection is required to reveal the single-stranded DNA tails necessary for strand invasion, and cyclin-dependent kinases appear to be required for its activation in S and G2 phase. End resection is robust in some specialized neurons, with one recent study in terminally differentiated rod cells demonstrating both NHEJ and HR activity (by its single-stranded annealing subpathway) upon a meganuclease-stimulated double-stranded break.^{54,55} This is encouraging for the application of targeted nucleases in other types of neurons, possibly in the central nervous system.

Other effector domains on DNA-binding proteins can also be used (Figure 4). TALEs and ZFPs can be engineered into designer transcription factors^{2,8,9,56,57} by attaching transcription modulation domains to the DNA binding modules. These designer transcription factors can be targeted to specific genes in the genome to modulate the activity of specific genes and probe their roles in brain function as well as disease states. It is important to consider that TALEs acting as transcription factors require physical binding to its DNA substrate to deliver its effector activity. As a result, the chromatin state of any locus of interest has to be taken into account. Open chromatin is typically defined as a region with DNase I hypersensitivity, so

careful study of genome-wide analyses for DNaseI hypersensitive tracts in the target cell type could be a way to improve the success of TAL effector targeting. Another open area is the fusion of TALEs to chromatin modifiers and regulators to modify the epigenomic state and affect transcriptional regulation. Other editing domains such as recombinases, transposases, or deaminases are alternative ways to modify genomic sequences. Finally, fluorescent proteins can be fused to DNA binding domains to illuminate chromosomal configuration.

Overall, the development of designer DNA binding proteins based on TALEs and ZFPs establishes a powerful platform for interrogating the function of genes in normal nervous system function as well as for engineering neuropsychiatric disease mechanisms. These technologies will be especially powerful when applied in specific cell type or brain circuits to dissect the role of gene mutations and epigenetic regulation in facilitate neural circuit function and computation. When combined with powerful readout methodologies at the molecular, circuit, and behavioral levels, targeted genome engineering technologies have the potential to significantly advance our understanding of the molecular basis of brain function.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: zhang_f@mit.edu. Mailing address: Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142.

Author Contributions

Both authors contributed equally to this work. P.D.H. and F.Z. performed literature review and wrote the manuscript.

Funding

Work described in this review was supported by grants from NIH (1R01NS073124), the McKnight, Gates, Damon Runyon, Keck, Klingenstein, and Simons Foundations, the Searle Scholars Program, Bob Metcalfe, and Mike Boylan. P.D.H. is supported by a James Mills Pierce Fellowship.

Notes

The authors declare no competing financial interest. Additional information can be found at the TAL Effector Resources Web site (<http://www.taleffectors.com>).

ACKNOWLEDGMENTS

We thank the entire Zhang laboratory for their support.

REFERENCES

- (1) Klug, A. (2005) Towards therapeutic applications of engineered zinc finger proteins. *FEBS Lett.* 579, 892–894.
- (2) Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G. M., and Arlotta, P. (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* 29, 149–153.
- (3) Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D., and Rebar, E. J. (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143–148.
- (4) Kim, Y. G., and Chandrasegaran, S. (1994) Chimeric restriction endonuclease. *Proc. Natl. Acad. Sci. U.S.A.* 91, 883–887.
- (5) Kim, Y. G., Cha, J., and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156–1160.
- (6) Perez, E. E., Wang, J., Miller, J. C., Jouvenot, Y., Kim, K. A., Liu, O., Wang, N., Lee, G., Bartsevich, V. V., Lee, Y. L., Guschin, D. Y., Rupniewski, I., Waite, A. J., Carpenito, C., Carroll, R. G., Orange, J. S., Urnov, F. D., Rebar, E. J., Ando, D., Gregory, P. D., Riley, J. L., Holmes, M. C., and June, C. H. (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* 26, 808–816.
- (7) Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S. Y., Bhagwat, A. S., Malani, N., Anguela, X. M., Sharma, R., Ivanciu, L., Murphy, S. L., Finn, J. D., Khazi, F. R., Zhou, S., Paschon, D. E., Rebar, E. J., Bushman, F. D., Gregory, P. D., Holmes, M. C., and High, K. A. (2011) In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* 475, 217–221.
- (8) Beerli, R. R., Dreier, B., and Barbas, C. F., 3rd. (2000) Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1495–1500.
- (9) Cong, L., Zhou, R., Kuo, Y.-c., Cunniff, M., and Zhang, F. (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.*, DOI: 10.1038/ncomms1962.
- (10) Rijkers, T., Peetz, A., and Ruther, U. (1994) Insertional mutagenesis in transgenic mice. *Transgenic Res.* 3, 203–215.
- (11) Nagy, A. (2000) Cre recombinase: the universal reagent for genome tailoring. *Genesis* 26, 99–109.
- (12) Capecchi, M. R. (1989) Altering the genome by homologous recombination. *Science* 244, 1288–1292.
- (13) Mansour, S. L., Thomas, K. R., and Capecchi, M. R. (1988) Disruption of the Proto-Oncogene Int-2 in Mouse Embryo-Derived Stem-Cells - a General Strategy for Targeting Mutations to Non-Selectable Genes. *Nature* 336, 348–352.
- (14) Rouet, P., Smih, F., and Jasin, M. (1994) Introduction of Double-Strand Breaks into the Genome of Mouse Cells by Expression of a Rare-Cutting Endonuclease. *Mol. Cell. Biol.* 14, 8096–8106.
- (15) Choulika, A., Perrin, A., Dujon, B., and Nicolas, J. F. (1995) Induction of Homologous Recombination in Mammalian Chromosomes by Using the I-SceI System of *Saccharomyces Cerevisiae*. *Mol. Cell. Biol.* 15, 1968–1973.
- (16) Paques, F., and Duchateau, P. (2007) Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr. Gene Ther.* 7, 49–66.
- (17) Pavletich, N. P., and Pabo, C. O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252, 809–817.
- (18) Bitinaite, J., Wah, D. A., Aggarwal, A. K., and Schildkraut, I. (1998) FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. U.S.A.* 95, 10570–10575.
- (19) Doyon, Y., Vo, T. D., Mendel, M. C., Greenberg, S. G., Wang, J., Xia, D. F., Miller, J. C., Urnov, F. D., Gregory, P. D., and Holmes, M. C. (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* 8, 74–79.
- (20) Carroll, D. (2011) Genome engineering with zinc-finger nucleases. *Genetics* 188, 773–782.
- (21) Davis, D., and Stokoe, D. (2010) Zinc Finger Nucleases as tools to understand and treat human diseases, *BMC Med.* 8, 42.
- (22) Ramirez, C. L., Foley, J. E., Wright, D. A., Muller-Lerch, F., Rahman, S. H., Cornu, T. I., Winfrey, R. J., Sander, J. D., Fu, F., Townsend, J. A., Cathomen, T., Voytas, D. F., and Joung, J. K. (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods* 5, 374–375.
- (23) Maeder, M. L., Thibodeau-Beganny, S., Osiaik, A., Wright, D. A., Anthony, R. M., Eichtinger, M., Jiang, T., Foley, J. E., Winfrey, R. J., Townsend, J. A., Unger-Wallace, E., Sander, J. D., Muller-Lerch, F., Fu, F., Pearlberg, J., Gobel, C., Dassie, J. P., Pruett-Miller, S. M., Porteus, M. H., Sgroi, D. C., Iafrate, A. J., Dobbs, D., McCray, P. B., Jr., Cathomen, T., Voytas, D. F., and Joung, J. K. (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell* 31, 294–301.
- (24) Sander, J. D., Dahlborg, E. J., Goodwin, M. J., Cade, L., Zhang, F., Cifuentes, D., Curtin, S. J., Blackburn, J. S., Thibodeau-Beganny, S., Qi, Y., Pierick, C. J., Hoffman, E., Maeder, M. L., Khayter, C., Reyon, D., Dobbs, D., Langenau, D. M., Stupar, R. M., Giraldez, A. J., Voytas, D. F., Peterson, R. T., Yeh, J. R., and Joung, J. K. (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat. Methods* 8, 67–69.
- (25) Segal, D. J. (2011) Zinc-finger nucleases transition to the CoDA. *Nat. Methods* 8, 53–55.
- (26) Gabriel, R., Lombardo, A., Arens, A., Miller, J. C., Genovese, P., Kaeppl, C., Nowrouzi, A., Bartholomae, C. C., Wang, J., Friedman, G., Holmes, M. C., Gregory, P. D., Glimm, H., Schmidt, M., Naldini, L., and von Kalle, C. (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* 29, 816–823.
- (27) Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y. L., Rupniewski, I., Beausejour, C. M., Waite, A. J., Wang, N. S., Kim, K. A., Gregory, P. D., Pabo, C. O., and Rebar, E. J. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* 25, 778–785.
- (28) Boch, J., and Bonas, U. (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* 48, 419–436.
- (29) Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512.
- (30) Moscou, M. J., and Bogdanove, A. J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501.
- (31) Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G., and Zhang, F. (2012) A transcription activator-like effector toolbox for genome engineering. *Nat. Protoc.* 7, 171–192.

- (32) Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J., and Voytas, D. F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39, e82.
- (33) Li, T., Huang, S., Zhao, X., Wright, D. A., Carpenter, S., Spalding, M. H., Weeks, D. P., and Yang, B. (2011) Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res.* 39, 6315–6325.
- (34) Morbitzer, R., Elsaesser, J., Hausner, J., and Lahaye, T. (2011) Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res.* 39, 5790–5799.
- (35) Weber, E., Gruetzner, R., Werner, S., Engler, C., and Marillonnet, S. (2011) Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One* 6, e19722.
- (36) Mahfouz, M. M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., and Zhu, J. K. (2011) De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2623–2628.
- (37) Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A. J., and Voytas, D. F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757–761.
- (38) Soldner, F., Laganieri, J., Cheng, A. W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L. I., Myers, R. H., Lindquist, S., Zhang, L., Guschin, D., Fong, L. K., Vu, B. J., Meng, X., Urnov, F. D., Rebar, E. J., Gregory, P. D., Zhang, H. S., and Jaenisch, R. (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 146, 318–331.
- (39) Carbery, I. D., Ji, D., Harrington, A., Brown, V., Weinstein, E. J., Liaw, L., and Cui, X. (2010) Targeted genome modification in mice using zinc-finger nucleases. *Genetics* 186, 451–459.
- (40) Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q. L., and Smith, A. (2008) Capture of authentic embryonic stem cells from rat blastocysts. *Cell* 135, 1287–1298.
- (41) Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R. E., Schulze, E. N., Song, H., Hsieh, C. L., Pera, M. F., and Ying, Q. L. (2008) Germline competent embryonic stem cells derived from rat blastocysts. *Cell* 135, 1299–1310.
- (42) Tong, C., Li, P., Wu, N. L., Yan, Y., and Ying, Q. L. (2010) Production of p53 gene knockout rats by homologous recombination in embryonic stem cells. *Nature* 467, 211–213.
- (43) Cui, X., Ji, D., Fisher, D. A., Wu, Y., Briner, D. M., and Weinstein, E. J. (2011) Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat. Biotechnol.* 29, 64–67.
- (44) Tesson, L., Usal, C., Menoret, S., Leung, E., Niles, B. J., Remy, S., Santiago, Y., Vincent, A. I., Meng, X., Zhang, L., Gregory, P. D., Aneqon, I., and Cost, G. J. (2011) Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol.* 29, 695–696.
- (45) Tachibana, M., Sparman, M., Ramsey, C., Ma, H., Lee, H. S., Penedo, M. C., and Mitalipov, S. (2012) Generation of chimeric rhesus monkeys. *Cell* 148, 285–295.
- (46) Yang, S. H., Cheng, P. H., Banta, H., Piotrowska-Nitsche, K., Yang, J. J., Cheng, E. C., Snyder, B., Larkin, K., Liu, J., Orkin, J., Fang, Z. H., Smith, Y., Bachevalier, J., Zola, S. M., Li, S. H., Li, X. J., and Chan, A. W. (2008) Towards a transgenic model of Huntington's disease in a non-human primate. *Nature* 453, 921–924.
- (47) Izsvak, Z., Frohlich, J., Grabundzija, I., Shirley, J. R., Powell, H. M., Chapman, K. M., Ivics, Z., and Hamra, F. K. (2010) Generating knockout rats by transposon mutagenesis in spermatogonial stem cells. *Nat. Methods* 7, 443–445.
- (48) Collin, J., and Lako, M. (2011) Concise review: putting a finger on stem cell biology: zinc finger nuclease-driven targeted genetic editing in human pluripotent stem cells. *Stem Cells* 29, 1021–1033.
- (49) Acland, G. M., Aguirre, G. D., Ray, J., Zhang, Q., Aleman, T. S., Cideciyan, A. V., Pearce-Kelling, S. E., Anand, V., Zeng, Y., Maguire, A. M., Jacobson, S. G., Hauswirth, W. W., and Bennett, J. (2001) Gene therapy restores vision in a canine model of childhood blindness. *Nat. Genet.* 28, 92–95.
- (50) Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C. C., Veres, G., Schmidt, M., Kutschera, I., Vidaud, M., Abel, U., Dal-Cortivo, L., Caccavelli, L., Mahlaoui, N., Kiermer, V., Mittelstaedt, D., Bellesme, C., Lahlou, N., Lefrere, F., Blanche, S., Audit, M., Payen, E., Leboulch, P., l'Homme, B., Bougneres, P., Von Kalle, C., Fischer, A., Cavazzana-Calvo, M., and Aubourg, P. (2009) Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 326, 818–823.
- (51) Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J. I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L. E., Wissler, M., Prinz, C., Rabbitts, T. H., Le Deist, F., Fischer, A., and Cavazzana-Calvo, M. (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302, 415–419.
- (52) Hanna, J., Wernig, M., Markoulaki, S., Sun, C. W., Meissner, A., Cassady, J. P., Beard, C., Brambrink, T., Wu, L. C., Townes, T. M., and Jaenisch, R. (2007) Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* 318, 1920–1923.
- (53) Hartlertode, A., Odate, S., Shim, I., Brown, J., and Scully, R. (2011) Cell cycle-dependent induction of homologous recombination by a tightly regulated I-SceI fusion protein. *PLoS One* 6, e16501.
- (54) Chan, F., Hauswirth, W. W., Wensel, T. G., and Wilson, J. H. (2011) Efficient mutagenesis of the rhodopsin gene in rod photoreceptor neurons in mice. *Nucleic Acids Res.* 39, 5955–5966.
- (55) Zan, Y., Haag, J. D., Chen, K. S., Shepel, L. A., Wigington, D., Wang, Y. R., Hu, R., Lopez-Guajardo, C. C., Brose, H. L., Porter, K. L., Leonard, R. A., Hitt, A. A., Schommer, S. L., Elegbede, A. F., and Gould, M. N. (2003) Production of knockout rats using ENU mutagenesis and a yeast-based screening assay. *Nat. Biotechnol.* 21, 645–651.
- (56) Stolzenburg, S., Rots, M. G., Beltran, A. S., Rivenbark, A. G., Yuan, X., Qian, H., Strahl, B. D., and Blancafart, P. (2012) Targeted silencing of the oncogenic transcription factor SOX2 in breast cancer. *Nucleic Acids Res.* DOI: 10.1093/nar/gks360.
- (57) Blancafart, P., Magnenat, L., and Barbas, C. F., 3rd (2003) Scanning the human genome with combinatorial transcription factor libraries. *Nat. Biotechnol.* 21, 269–274.
- (58) Anderson, J. L., Head, S. I., Rae, C., and Morley, J. W. (2002) Brain function in Duchenne muscular dystrophy. *Brain* 125, 4–13.