Kursad Turksen *Editor*

Genome Editing

Springer

Genome Editing

Kursad Turksen
Editor

Genome Editing

Springer

Editor
Kursad Turksen
Ottawa, ON, Canada

ISBN 978-3-319-34146-0 ISBN 978-3-319-34148-4 (eBook)
DOI 10.1007/978-3-319-34148-4

Library of Congress Control Number: 2016945569

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Contents

Contributors

Samvel Avagyan Life Sciences Solutions, Thermo Fisher Scientific, Carlsbad, CA, USA

CIRM Bridges Program, San Diego State University, San Diego, CA, USA

Rensselaer Polytechnic Institute, Troy, NY, USA

Alexander G. Bassuk Department of Pediatrics, The University of Iowa Children's Hospital, Iowa City, IA, USA

Channabasavaiah B. Gurumurthy Developmental Neuroscience, Munroe Meyer Institute for Genetics and Rehabilitation, University of Nebraska Medical Center, Omaha, NE, USA

Mouse Genome Engineering Core Facility, Office of the Vice Chancellor for Re-search, University of Nebraska Medical Center, Omaha, NE, USA

Izuho Hatada Laboratory of Genome Science, Biosignal Genome Resource Center, Institute for Molecular and Cellular Regulation, Gunma University, Maebashi, Gunma, Japan

Takuro Horii Laboratory of Genome Science, Biosignal Genome Resource Center, Institute for Molecular and Cellular Regulation, Gunma University, Maebashi, Gunma, Japan

Chun-Wei Hsu Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Sally Justus Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Harunobu Kagawa Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto, Japan

Ioannis Karakikes Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA

Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

Shin-Il Kim Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto, Japan

Nicole A.J. Krentz Diabetes Research Group, Child and Family Research Institute, Vancouver, BC, Canada

Departments of Surgery/Cellular and Physiological Sciences, University of British Columbia, Vancouver, BC, Canada

Uma Lakshmipathy Life Sciences Solutions, Thermo Fisher Scientific, Carlsbad, CA, USA

K.C. Kent Lloyd Mouse Biology Program, University of California, Davis, Davis, CA, USA

Francis C. Lynn Diabetes Research Group, Child and Family Research Institute, Vancouver, BC, Canada

Departments of Surgery/Cellular and Physiological Sciences, University of British Columbia, Vancouver, BC, Canada

Vinit B. Mahajan Department of Ophthalmology, Iowa University Hospitals and Clinics, Iowa City, IA, USA

Tomoji Mashimo Institute of Experimental Animal Sciences, Graduate School of Medicine, Osaka University, Osaka, Japan

Lluis Montoliu Department of Molecular and Cellular Biology, National Centre for Biotechnology (CNB-CSIC), Madrid, Spain

CIBERER-ISCIII, Madrid, Spain

Fabian Oceguera-Yanez Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto, Japan

Masato Ohtsuka Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, Isehara, Kanagawa, Japan

Stephane Pelletier Department of Immunology, St. Jude Children's Research Hospital, Memphis, TN, USA

Rolen M. Quadros Mouse Genome Engineering Core Facility, Office of the Vice Chancellor for Research, University of Nebraska Medical Center, Omaha, NE, USA

Masahiro Sato Section of Gene Expression Regulation, Frontier Science Research Center, Kagoshima University, Kagoshima, Japan

Timon Seeger Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA

Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

Davide Seruggia Department of Molecular and Cellular Biology, National Centre for Biotechnology (CNB-CSIC), Madrid, Spain

CIBERER-ISCIII, Madrid, Spain

Department of Hematology/Oncology, Boston Children's Hospital, Boston, MA, USA

Kyle Sylakowski Life Sciences Solutions, Thermo Fisher Scientific, Carlsbad, CA, USA

University of Pittsburgh, Pittsburgh, PA, USA

Vittavat Termglinchan Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA

Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

Yi-Ting Tsai Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Stephen H. Tsang Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Department of Pathology and Cell Biology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Edward S. Harkness Eye Institute, New York-Presbyterian Hospital/Columbia University Medical Center, New York, NY, USA

Knut Woltjen Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto, Japan

Hakubi Center for Advanced Research, Kyoto University, Kyoto, Japan

Wei-Pu Wu Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Wen-Hsuan Wu Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Joseph C. Wu Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA

Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA

Andrew Zheng Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Genome Editing with Targetable Nucleases

Stephane Pelletier

Abstract For decades, genome engineering relied on techniques that took years to master, required the generation of large and often complex DNA constructs containing selection markers, and could be applied to only a few organisms. However, the ease and efficiency of current technologies to edit genomes are unprecedented. With the advent of targetable nucleases, most notably the CRISPR-Cas9 technology, genomes of all species are now easily accessible to modifications. This advance has provided countless opportunities not only to further our understanding of gene functions and disease mechanisms but also to correct disease-causing mutations, modify crops and livestock, and perhaps modify our environment. This chapter discusses the advances in genome-editing technologies and their current and future applications.

Keywords Clustered regularly interspaced palindromic repeats (CRISPR)–Cas9 (CRISPR-associated protein 9) • CRISPR from *Prevotella* and *Francisella* 1 (Cpf1) • Genome editing • Genome engineering • Homologous recombination (HR) • Meganucleases • Nonhomologous end-joining (NHEJ) • Transcription activator-like effector nucleases (TALENs) • Zinc-finger nucleases (ZFNs)

Introduction

Technologies to manipulate genomes have greatly improved during the past few years. For decades, researchers have relied on homologous recombination (HR), a naturally occurring but infrequent event, to introduce or correct DNA alterations at specific genomic loci. Although the approach has provided invaluable insights into gene function and disease mechanisms, the low frequency of correct targeting events in many organisms and thus the need for large and complex DNA constructs as well as drug selection have limited its widespread use as a genome-editing strategy.

S. Pelletier, Ph.D. (✉)
Department of Immunology, St. Jude Children's Research Hospital,
Memphis, TN 38105, USA
e-mail: stephane.pelletier@stjude.org

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_1

Early studies on HR in eukaryotic cells showed that DNA double-strand ends, resulting from DNA double-strand breaks (DSBs), are highly recombinogenic [1]. This property of DNA double-strand ends prompted researchers to test whether the voluntary insertion of a DSB at a particular site within a genome would also promote HR at or near that site. Several groups have used a two-step strategy in which a meganuclease restriction site is inserted by conventional low-frequency HR and a DSB is introduced by expression of the cognate meganuclease. They found that voluntary insertion of DSBs at any given site promotes HR and gene replacement by several orders of magnitude [2–6]. However, the two-step strategy used in these studies did not allow for the rapid reprogramming of the nuclease and significantly limited the application of this strategy for genome editing.

In search of a more flexible reagent for introducing DNA DSBs within genomes, the Carroll group investigated the use of a new class of artificial endonucleases called the zinc-finger nucleases (ZFNs) that was developed by the Chandrasegaran group [7–9]. These nucleases function as dimers and use eukaryotic DNA-binding domains, the zinc-finger proteins (ZFPs) [10], to guide the catalytic domain of the nonspecific endonuclease Fok1 [11–14] to virtually any location within genomes. In contrast to meganucleases, which recognize predefined DNA sequences, ZFNs can be reprogrammed by modifying the ZFPs to target any desired location within genomes. Importantly, the use of ZFNs can improve site-specific gene replacement by HR by several orders of magnitude [9, 15]. However, the widespread adoption of ZFNs by the scientific community has been hampered by the complexity of their assembly and variable efficacy in vivo [16].

A decade after ZFNs were implemented as genome-editing tools, another class of targetable nucleases was developed. Similar to ZFNs, the transcription activator-like effector (TALE) endonucleases (TALENs) result from the fusion of engineered DNA-binding proteins (TALEs) and the catalytic domain of Fok1 [17, 18]. Similar to ZFNs, these enzymes function as dimers and efficiently cleave DNA to enable genome editing, and they can be reprogrammed to target virtually any location within genomes [19]. Although the cloning of TALE repeats has been challenging because of the high sequence homology between repeats, the simplicity and predictability of the TALE code (see following) enable rapid reprogramming of these endonucleases.

In 2012, the demonstration by Jinek and colleagues [20] that the CRISPR (clustered regularly interspaced short palindromic repeat)–Cas9 (CRISPR-associated protein 9) system from *Streptococcus pyogenes* can be adapted for genome engineering sparked a revolution in the field of genome editing. The CRISPR-Cas9 system for genome editing consists of a small RNA molecule, called single guide RNA (sgRNA), and a nonspecific endonuclease, Cas9 [20]. The small RNA molecule forms a complex with Cas9 and provides sequence specificity to the system by interacting with its genomic target according to the Watson–Crick base-pairing rules. In contrast to ZFNs and TALENs, which require complete reengineering of the enzymes for reprogramming, the simple modification of the sgRNA is sufficient to reprogram Cas9 to target any sequence of interest [20–22]. The simplicity of CRISPR-Cas9 reprogramming, its high efficacy, and its multiplexing capability

have propelled this technology to the forefront of genome editing. Can this technology be surpassed? A recent study identified a novel class 2 CRISPR effector system in *Francisella novicida* U112 having features that are slightly distinct from the CRISPR-Cas9 system and that can be readily used as a genome-editing tool [23].

This chapter describes the different classes of targetable nucleases and discusses their applications as genome-editing tools.

Zinc-Finger Nucleases

ZFNs are artificial chimeric endonucleases formed by the fusion of a versatile class of eukaryotic transcription factor DNA-binding domains, the zinc-finger proteins (ZFPs), and the catalytic domain of Fok1 (Fig. 1). In contrast to conventional type II endonucleases that cleave DNA within or close to their binding sites, Fok1 is a type IIS endonuclease which cleaves DNA distally from its binding site [11–14]. Both the DNA-binding domain and the nuclease activity can be physically separated [14, 24–26], allowing the fusion of the catalytic domain of Fok1 to other DNA-binding domains such as ZFP or TALEs. Fok1 needs to dimerize to exert catalytic activity [8]. The low affinity between the Fok1 catalytic subunits is insufficient to promote dimerization and prevents random DNA cleavage. The targeting of two Fok1 subunits to any given locus in the genome is guided by DNA-binding domains (ZFPs, TALEs, or Cas proteins; see following), which recognize two adjacent binding sites in opposite directions [8]. Appropriate spacing between the two binding sites is required to allow dimerization of the catalytic subunits (Fig. 1) [8, 19, 27].

In ZFNs, target specificity is provided by the ZFP domain. ZFPs are modular DNA-binding domains that were originally identified in eukaryotic transcription factors [28, 29]. ZFPs consist of tandem arrays of Cys2-His2 fingers, each binding a zinc (II) ion to form the DNA-binding domain. ZFPs bind DNA by inserting an α-helix into the major groove of the DNA double helix, and each finger recognizes approximately 3 bp of genomic DNA sequence [30, 31]. Three or more fingers are required to provide sufficient affinity for binding to genomic DNA, and ZFNs with three to six fingers are generally used for introducing DNA DSBs. Each array of three to six fingers recognizes 9 to 18 bp of genomic DNA, specifying 18 to 36 nucleotides for each pair of ZFNs. This arrangement provides remarkable specificity, especially considering that an 18-bp sequence is predicted to be found less than once in the human genome. Tandem repeat arrays of fingers bind DNA from the 3′-end to the 5′-end. The 3′-most triplet is recognized by the N-terminal finger, and so forth [30] (Fig. 1). To allow dimerization of the Fok1 catalytic subunits, opposing DNA sequences recognized by the pair of ZFNs must be separated by a 5- to 6-bp-long spacer [9, 32–34] (Fig. 1). One of the major drawbacks of using ZFNs for genome editing is that their binding is sensitive to the epigenetic status of the target DNA [35, 36].

Modular assembly methods and combinatorial selection methods have been developed to engineer specific ZFPs. Modular assembly methods are based on the functional autonomy of each finger to associate with the 3-bp DNA segments.

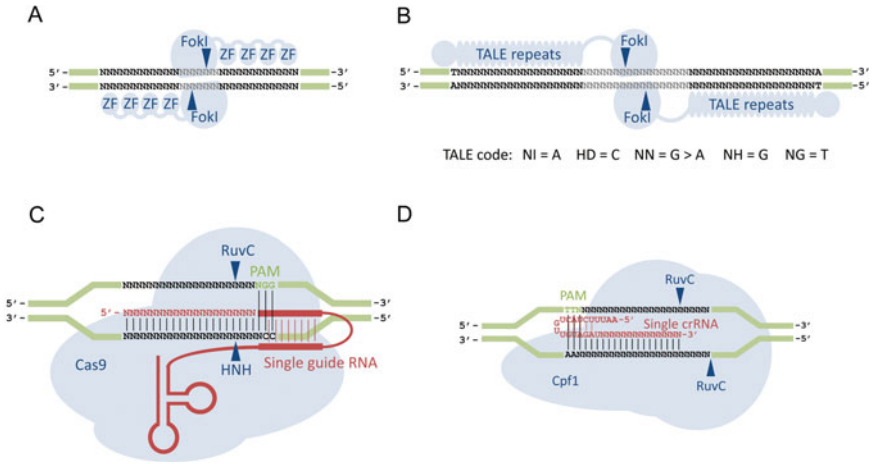

Fig. 1 Schematic representation of targetable nucleases for genome editing. (**a**) Schematic representation of a pair of zinc-finger nucleases (ZFN), each containing four zinc-finger (ZF) proteins, that recognizes a 12-nucleotide-long target sequence. A ZFN target site contains two zinc-finger-binding sites separated by a 5- to 7-nucleotide-long segment to enable FokI dimerization. ZFNs bind DNA "backward." The N-terminal ZFP recognizes the 3′-triplet of the recognition sequence and so forth. (**b**) Schematic representation of a pair of TALENs, each containing an array of 20 TALE repeats that defines a 20-nucleotide-long target sequence. A TALEN target site contains two TALE repeat binding sites separated by a 12- to 20-nucleotide-long segment accommodating for the requirement of extra TALE sequences at the C-terminus and FokI dimerization. Unlike ZFNs, which bind DNA 3′ to 5′, TALENs bind DNA 5′ to 3′ with the most N-terminal repeat recognizing the most 5′ nucleotide. The specificity of TALE repeats is provided by repeat variable di-residues at positions 12 and 13 of the 35-amino-acid TALE repeats. Four repeat variable di-residues NI, HD, NN, and NG, defining adenine (A) cytosine (C), guanine (G), and thymidine (T), respectively, are commonly used. In addition to the TALE repeats, additional protein sequences are required for TALENs to recognize their target DNA. At the N-terminus, two small repeats recognize a thymidine at position 0. (**c**) Schematic representation of the CRISPR-SpCas9 system for genome editing in complex with its target DNA. Cas9 is an endonuclease containing RuvC and HNH DNA endonuclease activity. The single guide RNA (*red*) is an artificial RNA molecule formed by the fusion of the crRNA with the tracrRNA by using a small RNA linker. The Cas9 endonuclease–RNA complex interacts with its target DNA sequence, the protospacer element, first by recognizing the small DNA element called the protospacer element adjacent motif located immediately downstream of the protospacer element. Alterations in the RuvC-like or HNH DNA nuclease activity can transform SpCas9 into a nickase that, when used as a pair, can reduce off-target activity. (**d**) Schematic representation of the CRISPR-FnCpf1 system for genome editing in complex with its target DNA. FnCpf1 is an endonuclease containing two RuvC-like endonuclease domains. Unlike SpCas9, which requires two small RNA molecules (or an artificial single guide RNA), FnCpf1 requires only a small crRNA molecule for activity. Like SpCas9, FnCpf1 also requires a PAM for activity. Interestingly, however, the PAM sequence is located immediately 5′ of the protospacer element, and Cpf1 cleaves its target DNA at the 3′-end of the protospacer element. Moreover, unlike SpCas9, which produces DNA breaks with blunt ends, FnCpf1 introduces scattered double-strand breaks, leaving 4- to 5-nucleotide-long 5′-overhangs

By linking three or more natural or artificial fingers together in a single polypeptide, a DNA-binding domain with sufficient affinity and specificity can be generated [37–41]. Although simple, these methods do not account for context-dependent factors that influence the affinity and specificity of ZFPs and have yielded ZFNs with poor efficacies, poor affinities, and high toxicities [42–44]. To account for context-dependent factors, pairs of fingers that function well together have been characterized and have been used for modular assembly [45–48]. Additional modifications, such as insertion of short or long linkers between fingers, can also be made to modular assembly methods to accommodate for B-form DNA or to skip base pairs [49]. Combinatorial selection methods involve the interrogation of large randomized libraries to select ZFPs with high DNA-binding affinities and high specificities [50, 51]. Although these strategies have yielded ZFNs with high activities and low toxicities [43, 44], only highly specialized laboratories have the expertise to perform these screens.

Several Web-based tools have been developed to help with the design and selection of ZFN pairs, such as the ZFN target site algorithm (http://mccb.umassmed.edu/ZFPmodularsearch.html); zinc finger tools (http://www.scripps.edu/barbas/zfdesign/zfdesignhome.php); ZiFit (http://zifit.partners.org/ZiFiT/); and ZFN-Site (http://ccg.vital-it.ch/tagger/targetsearch.html).

Transcription Activator-Like Effector Endonucleases

Similar to ZFNs, TALENs are artificial endonucleases that use TALE DNA-binding domains as DNA-binding units and the Fok1 endonuclease as the catalytic subunit. As mentioned previously, Fok1 functions as a dimer and therefore two TALE-Fok1 subunits must be brought together at any given locus for DNA cleavage. This process is mediated by the TALE DNA-binding domains, which are designed to recognize two adjacent and opposite DNA segments separated by 12 to 20 bp to allow Fok1 dimerization [19].

TALEs are transcription factors secreted by the gram-negative plant bacterium *Xanthomonas* during host invasion [52]. These transcription factors bind and regulate host gene expression to support bacterial virulence, proliferation, and dissemination. TALEs are composed of an N-terminal domain containing a type III translocation signal, a central region containing tandem repeats responsible for DNA binding of TALEs to plant DNA, and a C-terminal domain containing nuclear localization signals and a transcriptional activation domain [52]. The fusion of the Fok1 endonuclease at the trimmed C terminus of an N-terminally truncated TALE has provided the framework for generating TALENs [19] (Fig. 1).

Differing from ZFPs in which each finger recognizes approximately 3-bp segments of target DNA, each TALE repeat, composed of a 33- to 35-amino-acid segment, specifies 1 bp [53, 54]. TALE repeats are highly similar in sequence, and the specificity is conferred by a pair of adjacent amino acids (typically residues 12–13) within each repeat [55, 56]. These amino acids are referred to as “repeat variable

di-residue" (RVD) and specify the TALE repeat code. Four major repeats with hypervariable residues NI, HD, NN, and NG are commonly used, which recognize adenine (A), cytosine (C), guanine (G), and thymidine (T) [57–60], respectively (Fig. 1). Additional hypervariable residues with increased specificity have also been reported [19, 53, 59, 61] (Fig. 1). However, arrays containing some of these alternative RVDs can have less activity [59, 61].

Differing from ZFNs, which require no additional protein segments for activity, TALENs require, in addition to the TALE repeats, some protein sequences on each end of the repeats. Although several TALE frameworks have been used to develop TALENs as a genome-editing tool, they all appear to have similar properties. One important property is the presence of two small repeats similar in structure but distinct in sequence from the TALE repeats. These repeats recognize an obligate thymidine at position 0 of the TALE target DNA sequence [55] (Fig. 1); in other words, every TALEN target DNA sequence should start with a T. Although 10–12 TALE repeats are sufficient for binding, 15–21 TALE repeats are usually used for targeting, providing additional specificity and affinity to the TALEN. As seen with ZFNs, the DNA-cleavage activity of TALENs is affected by the epigenetic status of the target sites, in particular cytosine methylation in vivo [62]. Strategies to overcome this limitation have been developed, such as chemical inhibition of DNA methyltransferases or the use of an alternative TALE module in which the asparagine residue at position 12 is conserved and the residue at position 13 is omitted [62, 63].

The simplicity of the TALE code represents a major advancement over the more complex triplet DNA recognition system used by ZFNs. The cloning of several repeats, however, has been more challenging because of sequence similarity. To circumvent this limitation, several strategies such as solid-phase cloning [64, 65] and ligation-dependent [66] and ligation-independent [67] assembly strategies have been developed to construct TALE repeats with high specificity.

Several Web-based tools have been developed to help with the design and selection of TALE pairs, such as TAL Effector Nucleotide Targeter 2.0 (TALE-NT, https://tale-nt.cac.cornell.edu/); Scoring Algorithm for Predicting TALEN Activity (SAPTA, http://baolab.bme.gatech.edu/Research/BioinformaticTools/TAL_targeter.html); TAL effectors (http://www.genome-engineering.org/taleffectors/); E-TALEN (http://www.e-talen.org/E-TALEN/); CHOPCHOP (https://chopchop.rc.fas.harvard.edu/); TALEN designer (http://www.talen-design.de/); ZiFit (http://zifit.partners.org/ZiFiT/); and Mojo Hand (http://www.talendesign.org/).

RNA-Guided Nucleases

Bacteria and Archaea have evolved an adaptive immune system that captures DNA segments from invading phages or plasmids and integrates them within their genome for future use as a defense mechanism. These short DNA fragments, typically 20 to 50 nucleotides long, are integrated within their genomes in between spacer elements of similar length [68]. These sequences are termed as clustered regularly

interspaced palindromic repeats because they are incorporated in small clusters within the host genome. After a second infection by phages of the same family (or other invading DNAs), small RNAs generated from these clusters associate with a Cas endonuclease and direct the cleavage of the invading phage genome. Of the several distinct CRISPR-Cas systems present in Bacteria and Archaea [69, 70], the type II CRISPR-Cas system has been adapted for genome editing [20].

The type II CRISPR-Cas system has three components. The first component is a processed RNA transcript called CRISPR RNA (crRNA) which is produced from the CRISPR array and contains a single spacer element as well as part of the repeat element. The second component is a *trans*-encoded RNA transcript called trans-activating crRNA (TracrRNA), which has sequence complementarity to the repeat sequence of the crRNA transcript. The third component is the Cas9 endonuclease, which possesses two nuclease domains: a RuvC-like domain near the N-terminus of the protein and an HNH nuclease domain in the middle of the protein [20]. All three components form a ribonucleic complex that recognizes specific genomic DNA sequences, called protospacer elements, specified by the spacer element of the crRNA. The formation of the DNA–RNA hybrid between the crRNA and the protospacer element promotes cleavage of the target DNA [20]. The sequences receptive to the action of Cas endonuclease also require a protospacer-adjacent motif (PAM), which is a short DNA sequence located immediately downstream (3′) of the protospacer element [20, 71–73]. The *S. pyogenes* Cas9 (SpCas9) endonuclease, for example, requires a 5-NGG-3′ sequence (where N represents any nucleotide) for optimal activity [74]. A variant of this sequence (5′-NAG-3′) also confers limited activity, whereas any other combination of triplets does not confer activity [74]. The PAM requirements for several CRISPR-Cas systems have been described, some of which have been adapted for genome editing [73, 75–80] (Table 1). Moreover, by

Table 1 Protospacer adjacent motif requirements for CRISPR-Cas9 systems adapted for genome editing

Species	PAM	References
Streptococcus pyogenes	NGG	[74]
Streptococcus pyogenes (VQR variant)	NGAG	[81]
Streptococcus pyogenes (VRER variant)	NGCG	[81]
Streptococcus mutans	NGG	[79]
Staphylococcus aureus	NNGGGT	[81, 246]
	NNGAAT	
	NNGAGT	
Streptococcus thermophilus (CRISPR3)	NGGNG	[73, 76, 80]
Streptococcus thermophilus (CRISPR1)	NNAAAAW	[76]
Campylobacter jejuni	NNNNACA	[76]
Neisseria meningitidis	NNNNGATT	[77, 78]
Pasteurella multocida	GNNNCNNA	[76]
Francisella novicida	NG	[76]
Treponema denticola	NAAAAN	[75]

using structural information, selection-based direct evolution, and combinatorial design, SpCas9 has been modified to recognize alternative PAM sequences, which significantly extends the range of sequences amenable for genome editing [81] (Table 1). Collectively, Cas9 endonucleases from various species enable the coverage of virtually all nucleotides of genomes.

The most commonly used CRISPR-Cas9 system for genome editing comes from *S. pyogenes*. The three-component system has been simplified by linking the crRNA and tracrRNA into an sgRNA [20] (Fig. 1). The simple modification of the guide sequence, a 20-nucleotide sequence corresponding to the protospacer element, within the sgRNA molecule is sufficient to target SpCas9 to virtually any region of the genome (Fig. 1).

CRISPR-Cas9 systems offer several advantages over TALENs or ZFNs for genome editing. First, CRISPR-Cas9 systems rely on a single endonuclease that remains constant and thus does not require reengineering for reprogramming. Reprogramming of CRISPR-Cas9 systems requires only the design and generation of a new sgRNA. Second, the simplicity of the system offers opportunities for genome-wide applications and multiplexing. Transduction of multiple sgRNAs together with the Cas9 endonuclease allows the simultaneous targeting of multiple sites. CRISPR libraries are now publicly available, and genome-wide screens have identified genes essential for cell survival and drug resistance in vitro [82–89] as well as genes involved in tumor growth and metastasis in vivo [90]. Another advantage of the CRISPR-Cas9 technology over ZFNs and TALENS is its ability to cleave genomic DNA regardless of its epigenetic status [74, 91]. Finally, differing from ZFNs and TALENS, which cleave DNA distally to their binding site, SpCas9 cleaves DNA within its binding sites [20]. The resolution of the DNA DSB by homologous recombination (HR) or nonhomologous end-joining (NHEJ) modifies the target site and prevents further cleavage by Cas9. This property of CRISPR-Cas9 represents a significant asset for genome editing by HR.

Target specificity is crucial for the successful application of genome editing tools. Thus, sgRNA selection plays an important role in tailoring mutations. Several studies have shown that SpCas9 can tolerate mismatches, RNA bulges, and DNA bulges between the guide sequence of the crRNA and the target sequence [20, 21, 74, 92–95]. Although there are no simple and definitive rules for SpCas9 specificity, studies have shown that the number and the position of these mismatches relative to the PAM sequence are important. DNA–RNA duplex mismatches, including DNA and RNA bulges, located near the PAM sequence impair Cas9 activity, whereas mismatches located distally are better tolerated. Although single-nucleotide mismatches have little effect on SpCas9 activity, two or more mismatches, depending on their position, can significantly impair SpCas9 activity. Similarly, no simple and definitive rules governing on-target efficacy have been established. However, recent studies have suggested that the nucleotide immediately upstream of the PAM sequence may affect the efficacy of Cas9-mediated DNA DSBs [91, 96–98].

Several open-access websites have been developed to help with the selection of sgRNAs. These websites accept a wide range of inputs and suggest on-target locations on the basis of their uniqueness within a genome and generate a list of potential

off-target loci associated with each on-target sequence. These websites include CRISPR Design (http://crispr.mit.edu/); E-CRISP (http://www.e-crisp.org/E-CRISP/); ZiFit (http://zifit.partners.org/ZiFiT/); and Cas-OFFinder (http://www.rgenome.net/).

During the implementation of the CRISPR-Cas9 technology for mouse genome engineering, we developed our own sgRNA selection procedure by using Cas-Offinder [99]. Although our procedure may seem more tedious than other guide RNA selection methods, it provides a comprehensive list of all potential guide RNAs along with their potential off-target sites within genomes. Our strategy also takes into account observations made by several groups who have worked on defining sgRNA selectivity [74, 92, 94, 95, 100], including those showing that SpCas9 can tolerate several mismatches (up to 8) located at the 5′-end of the sgRNA [20]. Cas-Offinder allows the interrogation of not only several reference genomes but also nonreference vertebrate genomes, such as the non-obese diabetic and Friend leukemia virus, strain B mice. Importantly, our method allows the design of highly selective sgRNAs that exhibit little or no off-target activity in mice [99, 101].

More recently, another class 2 CRISPR-Cas system has been adapted for genome editing [23]. The type V CRISPR-Cas system from *Francisella novicida* (Fn) comprises a large protein called CRISPR from *Prevotella* and *Francisella* 1 (Cpf1) and a small crRNA. Unlike type II CRISPR-Cas systems, CRISPR-Cpf1 systems do not require tracrRNAs for function. The 5′-end of the crRNA contains a highly conserved spacer region that is predicted to form a short hairpin loop. Alteration of the sequence or the hairpin disrupts CRISPR-FnCpf1 cleavage activity, suggesting that FnCpf1 recognizes a combination of sequence-specific and structural features of the stem loop. The crRNA–Cpf1 complex cleaves DNA that is preceded by a short T-rich PAM sequence and introduces scattered DSBs with a 4- to 5-nucleotide-long 5′-overhang (Fig. 1). Similar to the CRISPR-Cas9 systems, the CRISPR-FnCpf1 endonuclease requires an 18-nucleotide-long spacer and can tolerate single-nucleotide mismatches between the guide sequence and the target DNA. In addition to FnCpf1, several other CRISPR-Cpf1 systems with distinct PAM requirements have been identified in bacteria (Table 2). Unlike CRISPR-Cas9 systems, however, and similar to TALENs and ZFNs, CRISPR-Cpf1 cleaves outside its recognition sequence—at nucleotide 18 on the nontargeted strand and nucleotide 23 on the targeted strand. This finding suggests that repeated cleavage might occur at the site until sufficient alterations are introduced to prevent recogni-

Table 2 Protospacer adjacent motif requirements for CRISPR-Cpf1 systems

Species	PAM	References
Franciselle novicida	TTN	[23]
Acidaminococcus sp. *BV3L6*	TTTN	[23]
Lachnospiraceae bacterium MA2020	TTTN	[23]
Moraxella bovoculi 237	(T/C)(T/C)N	[23]

tion of the target site by Cpf1. Nevertheless, FnCpf1 and related systems represent a great addition to our armory of genome editing tools.

Target Specificity

Target specificity is fundamental for the successful application of nuclease-based genome editing. Excessive cleavage of off-target loci can cause cytotoxicity and confound the interpretation of genetically modified organisms. In the context of clinical applications, off-target cleavage may have deleterious consequences if it modifies loci other than the ones intended. Thus, reagents need to be optimized, and potential off-target cleavage sites need to be identified.

Binding specificities of ZFPs and TALEs are usually determined in vitro by using the systematic evolution of ligands by exponential enrichment [102]. Information obtained from using this in vitro approach can then be used to interrogate the genome of interest and generate a list of potential off-target loci. These potential off-target loci can in turn be interrogated by direct sequencing of the loci [103, 104]. Binding specificity of CRISPR-Cas9 is governed by the Watson–Crick base-pairing rules. Potential off-targets are typically identified by using bioinformatics tools. Although whole-genome sequencing studies show that off-target mutagenesis is rare in mice and cultured stem cells edited by using CRISPR-Cas systems, the careful selection of sgRNAs is recommended [101, 105–109].

Two main approaches have been used to limit off-target cleavage by ZFNs and TALENs. The first approach consists of selecting ZFNs or TALENs with long DNA recognition sites (12–18 bp). As already mentioned, an 18-nucleotide sequence is found less than once in a 3-billion base-pair genome (approximately the size of the human genome). The second approach is to use obligate heterodimers of FokI that can prevent random dimerization of the nonspecific nuclease, and, consequently, random DNA cleavage [110]. Minimizing the off-target cleavage by ZFNs has also been achieved by reducing the half-life of ZFNs. Addition of an arginine residue at the N-terminus of ZFNs reduces toxicity in cells, possibly by preventing excessive off-target cleavage [44]. In addition to the careful selection of sgRNA, several other strategies have been proposed to minimize the cleavage of off-target loci by CRISPR-SpCas9; these include use of the SpCas9-D10A mutant, which inactivates the RuvC-like nuclease domain of SpCas9 [20], in combination with a pair of offset sgRNAs [111, 112]. This double nickases strategy can lead to a 100- to 1500-fold reduction of target activity in cell lines. Another strategy is to use shorter guide RNA molecules. Although counterintuitive, the use of this strategy reduces off-target cleavage by several thousand fold. In combination with paired nickases, this strategy can further improve selectivity [100, 113]. Another strategy is to combine the use of paired offset sgRNAs with catalytically inactive Cas9 fused to the catalytic subunit of the nonspecific endonuclease FokI. By using this strategy, off-target cleavage can be limited by

approximately sevenfold over that by the paired nickase strategy [114]. In this case, a spacer length of 13 to 17 bp is required for maximum activity. The target specificity of the newly identified CRISPR-Cpf1 systems is still unknown, and additional studies are required to better define their specificity.

Genome Editing with Targetable Nucleases

Fundamental to genome editing with targetable nucleases are the DNA-repair mechanisms involved in the resolution of DNA DSBs. In eukaryotic cells, introduction of a DNA DSB by targetable nucleases stimulates DNA-repair mechanisms, of which NHEJ and homology-directed repair (HDR or HR) are the most prominent [115]. Resolution of DNA DSBs by NHEJ promotes the ligation of DSB ends in an error-prone manner and often results in the small insertion or deletion (indels) of genetic material at the break site. In some cases, NHEJ can also result in intra- or interchromosomal translocations when two or more breaks are introduced. Resolution of DSBs by the HDR pathway is generally considered error free, and the sister chromatid is used as a template for repair. Taking advantage of this repair mechanism by providing a user-defined DNA repair template (donor template) allows the insertion of specific mutations, including large DNA elements.

Gene inactivation can be achieved by introducing a single DNA DSB within an exon downstream of the translational start site (Fig. 2a). The resolution of DSBs by NHEJ can introduce several genetic modifications such as nonsense, missense, and frameshift mutations, which often result in gene inactivation by engaging the nonsense-mediated decay pathway [116]. This strategy is commonly used for genome-wide CRISPR-mediated screens [82–89] and gene inactivation in vivo [99, 107, 109, 117].

Gene correction and site-directed mutagenesis can be achieved by introducing a single DNA DSB near the location of the intended correction or mutation (Fig. 2b). Resolution of DSBs by HR is promoted by providing user-designed DNA templates that have sequence homology on both sides of the desired mutation(s). This strategy has been used to insert point mutations and large DNA fragments in vitro and in model organisms (Fig. 2c) [15, 99, 101, 107, 117–123].

Gene inactivation can also be achieved by introducing two DSBs flanking the gene of interest (Fig. 2d). Resolution of these DSBs by NHEJ can result in exclusion of the intervening region [99, 117, 118, 124–129]. However, resolution of these breaks can also result in insertion or deletion of small DNA segments at both loci and might not generate null alleles.

The introduction of two DNA DSBs within the introns of genes can be used to insert recombinase recognition sites (e.g., loxP or Flp sites) to generate conditional alleles (Fig. 2d). DNA fragments containing loxP sites flanked by regions of homology to the target sites can serve as donor DNA for HR [99, 107, 117]. As already described, the resolution of DSBs by NHEJ can also result in deletion of the intervening region, thereby creating a null allele. Often in an attempt to generate mice bearing a conditional allele, null alleles are also generated by NHEJ [99]. Typically, targeting strategies that can generate both alleles simultaneously are designed [99].

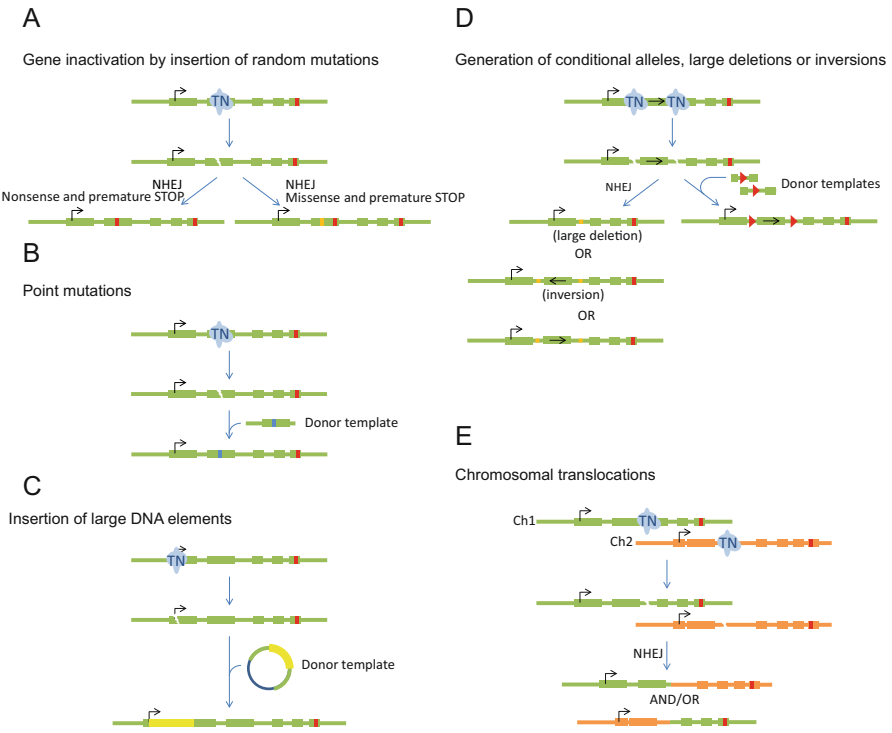

Fig. 2 Targeting strategies used to introduce mutations by using targetable nucleases. (**a**) Gene inactivation and insertion of random mutations can be achieved by the use of a single targetable nuclease. Resolution of the double-strand breaks (DSBs) by NHEJ generates various alterations at the locus, such as nonsense, missense, and frameshift mutations, as well as insertion or deletion of large DNA fragments. (**b**) Specific point mutations can be introduced in genomes by using a single targetable nuclease. Resolution of the single DSB by homologous recombination (HR) by using a donor template containing the specified mutations and regions of homology flanking the mutation(s) generates allele(s) with the desired mutation. (**c**) Insertion of large DNA elements can be achieved by introducing a single DNA DSB by using a single targetable nuclease. Resolution of the DSB by HR by using a donor template containing the desired DNA sequence to be inserted (e.g., YFP open reading frame) flanked by a sequence of homology corresponding to the target locus generates one or more alleles with the desired insertion. (**d**) Generation of conditional alleles and large deletions can be achieved by using two targetable nucleases designed to target introns flanking one or more critical exons of a gene. The resolution of these breaks by nonhomologous end-joining (NHEJ) can result in the exclusion or inversion of the intervening region, likely generating null alleles. Resolution of the DNA DSBs by NHEJ can also result in the generation of indels at the break sites. Resolution of the breaks by HR in the presence of donor templates containing recombinase recognition sequences (e.g., loxP sites) flanked by homology arms corresponding to the target sites generates floxed alleles. (**e**) Chromosomal translocations can be generated by introducing DNA DSBs by using two targetable nucleases aiming at distinct chromosomal regions where the translocation is desired. Resolution of these breaks by NHEJ can result in the translocation of the two chromosomes. Insertion of loxP sites on each chromosome could also be introduced to make the translocation inducible

Insertion of DSBs on separate chromosomes can also promote chromosomal translocations [130–136]; this is of particular interest in the case of cancer, wherein translocations often result in cellular transformation, and mimicking these translocations in cells or mice can provide critical insights into disease pathogenesis. Homology-directed repair templates containing segments homologous to both chromosomes can be used to promote chromosomal translocation. Alternatively, loxP sites can be inserted within both chromosomes and translocation can be promoted by Cre-mediated recombination [99, 137].

Targetable nucleases can also be used to inactivate gene families with high sequence homology. A single pair of ZFNs or TALENs or a single sgRNA can be designed to target several family members [99] (Fig. 2a).

Another application of targetable nucleases is insertion of the DNA transgene into genomic safe harbors by using nonintegrating viral vectors. The PPP1R12C locus on the human chromosome 19 is a transcriptionally competent region in which transgenes can be safely inserted without inducing adverse effects. This region is also known as AAVS1, and the insertion of DSBs by targetable nucleases can promote the insertion of large DNA elements flanked by regions of homology to the AAVS1 locus [138–140].

Promoting Homologous Recombination over Nonhomologous End-Joining

DSB repair pathways are in constant competition with one another. Because of this, the introduction of precise mutations by HR using targetable nucleases continues to be a major challenge. Two main strategies have been proposed to circumvent this problem. The first strategy consists of inhibiting NHEJ, the most important DSB repair mechanism in eukaryotic cells. The second strategy is to deliver targetable nucleases during the late S- and G_2-phases of the cell cycle when the HR pathway is most active.

Studies on DSB resolution pathways have identified several factors involved in the canonical NHEJ pathway [115]. These factors include the heterodimers Ku70/Ku80, which bind DNA DSBs and function as scaffolds to recruit the NHEJ machinery. Another important factor in NHEJ is DNA ligase IV, which is essential for DNA end-joining. Inhibition of NHEJ by inhibiting the expression of the Ku70/Ku80 heterodimers or DNA ligase IV as well as inhibition of DNA ligase IV binding to DNA DSB by the small molecule SCR7 can improve HR by several fold while reducing NHEJ both in cell lines and in mouse embryos [141–143].

NHEJ operates throughout the cell cycle whereas HR is restricted to the late S- and G_2-phases. Thus, delivering targetable nucleases at the late S- and G_2-phases can likely promote HR over NHEJ. Indeed, the delivery of CRISPR-Cas9 reagents into cells treated with nocodazole, which causes cell-cycle arrest at the G_2/M-phase, leads to a dramatic increase in HR (up to 38 %) relative to untreated cells [144].

Delivery Methods

The delivery of targetable nucleases and donor DNA for HR into cells is also essential for the successful application of genome editing. Delivery of DNA plasmids that encode targetable nucleases and donor templates by using electroporation or transfection using chemical reagents is common in transformed cells as well as select embryonic stem cells. However, these methods are usually inefficient in primary cell cultures and may induce toxicity. Several delivery methods have been developed to overcome these limitations.

One approach being investigated is the delivery of in vitro translated targetable nucleases by using reagent-free transfection strategies. These strategies rely on the chemical or genetic modification of enzymes to make them membrane permeable. For example, conjugation of poly-arginine peptides to TALENs favors their uptake by cells, and the reversible modification enables TALEN-mediated gene knockout at a rate similar to that by plasmid transfection [145]. The genetic fusion of TALEN to the cell-permeable TAT peptide also enables uptake of the fusion protein by cells while maintaining activity [146]. The genetic fusion of ZFNs to transferrin enables the uptake of the ZFN fusion protein and leads to site-specific in situ cleavage of the target locus [147]. Efficient gene disruption in cultured cells can also be achieved by using cell-penetrating peptide-conjugated Cas9 and sgRNA [148]. The positively charged nature of ZFPs makes ZFNs membrane permeable, and internalization of unconjugated ZFNs leads to efficient endogenous gene disruption in mammalian cells [149]. Another chemical-free delivery method under investigation is electroporation of in vitro-translated and in vitro-assembled targetable nucleases, in particular sgRNA–Cas9 complexes. This delivery method is less toxic than, but as efficient as, DNA electroporation for introducing DNA DSBs [150]. Chemical reagents-dependent transfection procedures have also been developed to deliver in vitro-assembled CRISPR-Cas9 complexes. This delivery method can increase the efficacy of CRISPR-Cas9-mediated genome editing by protecting the nuclease complex from being neutralized by serum proteins blood cells and the extracellular matrix or from being degraded by the endosomal/lysosomal pathway [151]. Electroporation, chemical transfection, or microinjection of mRNA transcripts encoding targetable nucleases is another widely used approach for genome editing in cell lines and zygotes [99, 101, 107, 109, 117, 152]. The limited off-target activity associated with these hit-and-run strategies is likely caused by the short half-lives of the in vitro-translated proteins or in vitro-transcribed mRNAs, which exert their activity for a shorter period of time than do plasmid DNA molecules and/or viral vectors (see following). Moreover, the random integration of plasmid DNA or viral vectors within genomes not only can affect the length of nuclease expression but also can interfere with genes in which the DNA has been inserted. In this regard, episomal viral vectors are preferred to integrating viruses.

Several viral vectors are currently being developed for targetable nuclease technology, such as adenovirus (AV), adeno-associated virus (AAV), lentivirus (LV), integrase-deficient lentivirus (IDLV), and baculovirus (BV) vectors. These vectors

have been extensively used to deliver targetable nucleases in vitro and in model organisms. Some of these delivery systems, however, have their limitations and might not be suitable for all targetable nucleases.

For example, LV and IDLV vectors have been used to deliver ZFNs and TALENS as well as donor plasmids in vitro [153–155]. However, their susceptibility to epigenetic modifications and silencing limits their efficacy. Moreover, the expression of intact TALENs can be severely impaired because of the extensive deletion of TALE repeats caused by switching of the reverse transcriptase template within TALE repeats [156, 157]. This limitation, however, can be surmounted by using alternative coding sequences that limit reverse transcriptase template switching [157] or by using reverse transcriptase-deficient viruses [158].

BV vectors have been used to deliver ZFNs and TALENs with high efficacy in embryonic stem cell cultures, and the large packaging capacity of BV vectors has enabled the packaging of ZFP and TALEN dimers into a single viral particle [159–162]. However, under certain propagation conditions, TALE repeats can also undergo rearrangement, which makes BV vectors less suitable for gene therapy and gene targeting in general [156, 163].

AV vectors represent an excellent platform for the delivery of targetable nucleases, because they can infect dividing and nondividing cells, accommodate large DNA fragments and code for proteins without the need for integrating host genomes. Moreover, AV vectors are one of the most prominent vectors used in gene therapy clinical trials [164]. AV vectors have been used to deliver ZFNs, TALENs, and RNA-guided nucleases both in vitro and in vivo [156, 165–170]. AV particles are large enough to accommodate dimers of ZFNs and TALENs, the large endonuclease Cas9 and sgRNA, and donor templates.

Nonpathogenic AAV vectors have been developed for efficient gene delivery and have yielded promising results in phase I to phase III clinical trials [171]. AAVs are replication-deficient viral particles that rely on unrelated viruses for replication. Their genetic material is composed of a 4.7-kb single-stranded DNA with two open reading frames. AAVs have been used to deliver ZFNs [172–174], and, despite their smaller packaging capacity, CRISPR-Cas9 systems [175–177].

Applications

Targetable nucleases have enabled genome editing in a plethora of mammalian and nonmammalian cells as well as model organisms commonly used in biomedical research laboratories, such as yeast [178, 179], *Caenorhabditis elegans* [180–182], *Drosophila* [183–185], zebrafish [61, 186–190], *Xenopus* [191–194], mouse [99, 101, 107, 109, 117, 195–199], rat [200–203], rabbit [204–207], and monkey [208–210]. In laboratory models, targetable nucleases have been used to determine gene function, study structure–function relationships, and create animal models of human diseases.

Targetable nucleases have also been used to modify crops [211–215], fungi [216, 217], and livestock [124, 218–221]. In plants, for example, targetable nucleases have been used to insert important traits such as disease resistance [214] and herbicide resistance [212, 213]. Gene inactivation in livestock has enabled the refinement of traits for xenotransplantation [222–225], and it has been used to generate animal models for human diseases [226]. Gene inactivation or insertion can also confer pathogen resistance [218, 227], increase productivity [220, 227, 228], and perhaps reduce the transmission of zoonotic disease.

Targetable nucleases also hold great promise for clinical applications in humans. Targetable nucleases can be used to correct and cure monogenic diseases, and the multiplexing capability of CRISPR-Cas9 can be used to correct and cure complex genetic disorders [21, 92]. Targetable nucleases can also be used for the targeted inactivation of genes associated with a certain trait or to inactivate viral infections. As proof of concept, targetable nucleases have been used to correct disease-causing mutations associated with sickle cell disease [229–231], hemophilia A and B [232–235], α1-antitrypsin deficiency [236], X-linked severe combined immunodeficiency in human hematopoietic stem cells [153], Duchenne muscular dystrophy in human myocytes [237, 238], Wiskott–Aldrich syndrome in hematopoietic stem cells [239], and cystic fibrosis in epithelial cells generated from patient-induced pluripotent stem cells from patients [240]. In vivo, targetable nucleases have been used to prevent hypercholesterolemia [241] and Duchenne muscular dystrophy [242].

Targetable nucleases have also been used for the targeted inactivation of the HIV1 co-receptor C–C chemokine receptor type 5 (CCR5) in T cells and hematopoietic stem cells [167, 243], and clinical trials on ZFN nucleases targeting CCR5 are currently underway [165]. Targetable nucleases not only offer an opportunity to target and inactivate genes involved in viral infection but also allow the inactivation of integrated viral vectors. A recent study used CRISPR-Cas9 technology to target the HIV1 LTR U3 region and efficiently excised the 9.7-kb fragment containing the integrated HIV genome [244]. Similar strategies have been employed to eliminate other latent and potentially pathogenic viruses from human genomes, such as the Epstein–Barr virus [245].

Conclusion

The development of targetable nucleases capable of introducing DNA DSBs to specific sites within genomes has greatly improved our ability to manipulate genomes and holds great promises for gene therapy. To achieve the full potential of the technology, however, several limitations inherent to the technologies themselves such as target specificity and delivery of the targetable nucleases and donor templates, as well as those associated with the DNA-repair mechanisms, need to be improved. Although improving target specificity and delivery has been the main focus of recent studies, developing a better understanding of the DNA-repair mechanisms and, more importantly, identifying ways to promote HR represents, in my opinion,

the most challenging and pressing issue. Inhibition of DNA ligase IV is already showing promising results in cell cultures and in model organisms and may be a translatable approach for in vivo gene therapy in humans.

Acknowledgments The author thanks Dr. Vani Shanker for editing the manuscript.

Funding The author received support from the American Lebanese Syrian Associated Charities.

References

1. Bollag RJ, Waldman AS, Liskay RM. Homologous recombination in mammalian cells. Annu Rev Genet. 1989;23:199–225.
2. Cohen-Tannoudji M, Robine S, Choulika A, Pinto D, El Marjou F, Babinet C, et al. I-SceI-induced gene replacement at a natural locus in embryonic stem cells. Mol Cell Biol. 1998;18(3):1444–8. Pubmed Central PMCID: 108858.
3. Donoho G, Jasin M, Berg P. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol Cell Biol. 1998;18(7):4070–8. Pubmed Central PMCID: 108991.
4. Gloor GB, Nassif NA, Johnson-Schlitz DM, Preston CR, Engels WR. Targeted gene replacement in *Drosophila* via P element-induced gap repair. Science. 1991;253(5024):1110–7.
5. Smih F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic Acids Res. 1995;23(24):5012–9. Pubmed Central PMCID: 307507.
6. Choulika A, Perrin A, Dujon B, Nicolas JF. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. Mol Cell Biol. 1995;15(4):1968–73. Pubmed Central PMCID: 230423.
7. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A. 1996;93(3):1156–60. Pubmed Central PMCID: 40048.
8. Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res. 2000;28(17):3361–9. Pubmed Central PMCID: 110700.
9. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, et al. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol. 2001;21(1):289–97. Pubmed Central PMCID: 88802.
10. Berg JM. Potential metal-binding domains in nucleic acid binding proteins. Science. 1986;232(4749):485–7.
11. Nwankwo D, Wilson G. Cloning of two type II methylase genes that recognise asymmetric nucleotide sequences: FokI and HgaI. Mol Gen Genet. 1987;209(3):570–4.
12. Looney MC, Moran LS, Jack WE, Feehery GR, Benner JS, Slatko BE, et al. Nucleotide sequence of the FokI restriction-modification system: separate strand-specificity domains in the methyltransferase. Gene. 1989;80(2):193–208.
13. Kita K, Kotani H, Sugisaki H, Takanami M. The FokI restriction-modification system. I. Organization and nucleotide sequences of the restriction and modification genes. J Biol Chem. 1989;264(10):5751–6.
14. Li L, Wu LP, Chandrasegaran S. Functional domains in Fok I restriction endonuclease. Proc Natl Acad Sci U S A. 1992;89(10):4275–9. Pubmed Central PMCID: 49064.
15. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. Science. 2003;300(5620):5764.
16. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 2010;11(9):636–46.

17. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 2010;186(2):757–61. Pubmed Central PMCID: 2942870.
18. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res. 2011;39(1):359–72. Pubmed Central PMCID: 3017587.
19. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011;29(2):143–8.
20. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337(6096):816–21.
21. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823–6. Pubmed Central PMCID: 3712628.
22. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. eLife. 2013;2, e00471. Pubmed Central PMCID: 3557905.
23. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. Cell. 2015;163:759–71.
24. Li L, Wu LP, Clarke R, Chandrasegaran S. C-terminal deletion mutants of the FokI restriction endonuclease. Gene. 1993;133(1):79–84.
25. Li L, Chandrasegaran S. Alteration of the cleavage distance of Fok I restriction endonuclease by insertion mutagenesis. Proc Natl Acad Sci U S A. 1993;90(7):2764–8. Pubmed Central PMCID: 46176.
26. Kim YG, Li L, Chandrasegaran S. Insertion and deletion mutants of FokI restriction endonuclease. J Biol Chem. 1994;269(50):31978–82.
27. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol. 2014;32(6):569–76. Pubmed Central PMCID: 4090141.
28. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. 1985;4(6):1609–14. Pubmed Central PMCID: 554390.
29. Brown RS, Sander C, Argos P. The primary structure of transcription factor TFIIIA has 12 consecutive repeats. FEBS Lett. 1985;186(2):271–4.
30. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science. 1991;252(5007):5809–17.
31. Berg JM. Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc Natl Acad Sci U S A. 1988;85(1):99–102. Pubmed Central PMCID: 279490.
32. Handel EM, Alwin S, Cathomen T. Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. Mol Ther. 2009;17(1):104–11. Pubmed Central PMCID: 2834978.
33. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature. 2005;435(7042):646–51.
34. Shimizu Y, Bhakta MS, Segal DJ. Restricted spacer tolerance of a zinc finger nuclease with a six amino acid linker. Bioorg Med Chem Lett. 2009;19(14):3970–2. Pubmed Central PMCID: 2709702.
35. Liu PQ, Rebar EJ, Zhang L, Liu Q, Jamieson AC, Liang Y, et al. Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. J Biol Chem. 2001;276(14):11323–34.
36. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell. 2008;31(2):294–301. Pubmed Central PMCID: 2535758.

37. Beerli RR, Barbas III CF. Engineering polydactyl zinc-finger transcription factors. Nat Biotechnol. 2002;20(2):135–41.
38. Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS, et al. Human zinc fingers as building blocks in the construction of artificial transcription factors. Nat Biotechnol. 2003;21(3):275–80.
39. Liu Q, Xia Z, Zhong X, Case CC. Validated zinc finger protein designs for all 16 GNN DNA triplet targets. J Biol Chem. 2002;277(6):3850–6.
40. Mandell JG, Barbas CF III (2006) Zinc finger tools: custom DNA-binding domains for transcription factors and nucleases. Nucleic Acids Res 2006;34(Web Server issue):W516–W523
41. Segal DJ, Beerli RR, Blancafort P, Dreier B, Effertz K, Huber A, et al. Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. Biochemistry. 2003;42(7):2137–48.
42. Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods. 2008;5(5):374–5.
43. Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, et al. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. Mol Ther. 2008;16(2):352–8.
44. Pruett-Miller SM, Connelly JP, Maeder ML, Joung JK, Porteus MH. Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. Mol Ther. 2008;16(4):707–17.
45. Gupta A, Christensen RG, Rayla AL, Lakshmanan A, Stormo GD, Wolfe SA. An optimized two-finger archive for ZFN-mediated gene targeting. Nat Methods. 2012;9(6):588–90. Pubmed Central PMCID: 3443678.
46. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods. 2011;8(1):67–9. Pubmed Central PMCID: 3018472.
47. Zhu C, Gupta A, Hall VL, Rayla AL, Christensen RG, Dake B, et al. Using defined finger-finger interfaces as units of assembly for constructing zinc-finger nucleases. Nucleic Acids Res. 2013;41(4):2455–65. Pubmed Central PMCID: 3575815.
48. Moore M, Klug A, Choo Y. Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. Proc Natl Acad Sci U S A. 2001;98(4):1437–41. Pubmed Central PMCID: 29275.
49. Moore M, Choo Y, Klug A. Design of polyzinc finger peptides with structured linkers. Proc Natl Acad Sci U S A. 2001;98(4):1432–6. Pubmed Central PMCID: 29274.
50. Greisman HA, Pabo CO. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. Science. 1997;275(5300):657–61.
51. Hurt JA, Thibodeau SA, Hirsh AS, Pabo CO, Joung JK. Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. Proc Natl Acad Sci U S A. 2003;100(21):12271–6. Pubmed Central PMCID: 218748.
52. Boch J, Bonas U. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol. 2010;48:419–36.
53. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009;326(5959):1501.
54. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009;326(5959):1509–12.
55. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science. 2012;335(6069):716–9. Pubmed Central PMCID: 3427646.
56. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, et al. Structural basis for sequence-specific recognition of DNA by TAL effectors. Science. 2012;335(6069):720–3. Pubmed Central PMCID: 3586824.
57. Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, et al. Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS One. 2012;7(9), e45383. Pubmed Central PMCID: 3454392.

58. Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun. 2012;3:968. Pubmed Central PMCID: 3556390.
59. Streubel J, Blucher C, Landgraf A, Boch J. TAL effector RVD specificities and efficiencies. Nat Biotechnol. 2012;30(7):593–5.
60. Meckler JF, Bhakta MS, Kim MS, Ovadia R, Habrian CH, Zykovich A, et al. Quantitative analysis of TALE-DNA interactions suggests polarity effects. Nucleic Acids Res. 2013;41(7):4118–28. Pubmed Central PMCID: 3627578.
61. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol. 2011;29(8):699–700.
62. Bultmann S, Morbitzer R, Schmidt CS, Thanisch K, Spada F, Elsaesser J, et al. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res. 2012;40(12):5368–77. Pubmed Central PMCID: 3384321.
63. Valton J, Dupuy A, Daboussi F, Thomas S, Marechal A, Macmaster R, et al. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem. 2012;287(46):38427–32. Pubmed Central PMCID: 3493886.
64. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol. 2012;30(5):460–5. Pubmed Central PMCID: 3558947.
65. Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, et al. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res. 2012;40(15), e117. Pubmed Central PMCID: 3424587.
66. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011;39(12), e82. Pubmed Central PMCID: 3130291.
67. Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol. 2013;31(1):76–81. Pubmed Central PMCID: 4142318.
68. Westra ER, Swarts DC, Staals RH, Jore MM, Brouns SJ, van der Oost J. The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annu Rev Genet. 2012; 46:311–39.
69. Makarova KS, Koonin EV. Annotation and classification of CRISPR-Cas systems. Methods Mol Biol. 2015;1311:47–75.
70. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol. 2011;9(6):467–77. Pubmed Central PMCID: 3380444.
71. Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468(7320):67–71.
72. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009;155(Pt 3):733–40.
73. Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J Bacteriol. 2008;190(4):1390–400. Pubmed Central PMCID: 2238228.
74. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31(9):827–32. Pubmed Central PMCID: 3969858.
75. Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat Methods. 2013;10(11):1116–21. Pubmed Central PMCID: 3844869.
76. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lecrivain AL, Bzdrenga J, et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous

type II CRISPR-Cas systems. Nucleic Acids Res. 2014;42(4):2577–90. Pubmed Central PMCID: 3936727.
77. Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. Proc Natl Acad Sci U S A. 2013;110(39):15644–9. Pubmed Central PMCID: 3785731.
78. Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C, et al. Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. Mol Cell. 2013;50(4):488–503. Pubmed Central PMCID: 3694421.
79. van der Ploeg JR. Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. Microbiology. 2009;155(Pt 6):1966–76.
80. Horvath P, Romero DA, Coute-Monvoisin AC, Richards M, Deveau H, Moineau S, et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. J Bacteriol. 2008;190(4):1401–12. Pubmed Central PMCID: 2238196.
81. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 2015;523(7561):481–5. Pubmed Central PMCID: 4540238.
82. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014;343(6166):84–7. Pubmed Central PMCID: 4089965.
83. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014;11(8):783–4. Pubmed Central PMCID: 4486245.
84. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014;343(6166):80–4. Pubmed Central PMCID: 3972032.
85. Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol. 2014;32(3):267–73.
86. Zhou Y, Zhu S, Cai C, Yuan P, Li C, Huang Y, et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. Nature. 2014;509(7501):487–91.
87. Parnas O, Jovanovic M, Eisenhaure TM, Herbst RH, Dixit A, Ye CJ, et al. A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. Cell. 2015;16:675–86.
88. Ma H, Dang Y, Wu Y, Jia G, Anaya E, Zhang J, et al. A CRISPR-based screen identifies genes essential for West-Nile-virus-induced cell death. Cell Rep. 2015;28:673–83.
89. Bassett AR, Kong L, Liu JL. A genome-wide CRISPR library for high-throughput genetic screening in *Drosophila* cells. J Genet Genomics. 2015;42(6):301–9.
90. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell. 2015;160(6):1246–60. Pubmed Central PMCID: 4380877.
91. Chari R, Mali P, Moosburner M, Church GM. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. Nat Methods. 2015;12:823–6.
92. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819–23. Pubmed Central PMCID: 3795411.
93. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol. 2013;31(9):822–6. Pubmed Central PMCID: 3773023.
94. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol. 2013;31(3):233–9. Pubmed Central PMCID: 3748948.
95. Lin Y, Cradick TJ, Brown MT, Deshmukh H, Ranjan P, Sarode N, et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Res. 2014;42(11):7473–85. Pubmed Central PMCID: 4066799.
96. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol. 2014;32(12):1262–7. Pubmed Central PMCID: 4262738.

97. Gagnon JA, Valen E, Thyme SB, Huang P, Akhmetova L, Pauli A, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS One. 2014;9(5), e98186. Pubmed Central PMCID: 4038517.
98. Xu H, Xiao T, Chen CH, Li W, Meyer CA, Wu Q, et al. Sequence determinants of improved CRISPR sgRNA design. Genome Res. 2015;25(8):1147–57. Pubmed Central PMCID: 4509999.
99. Pelletier S, Gingras S, Green DR. Mouse genome engineering via CRISPR-Cas9 for study of immune function. Immunity. 2015;42(1):18–27.
100. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol. 2014;32(3):279–84. Pubmed Central PMCID: 3988262.
101. Martinez J, Malireddi RK, Lu Q, Cunha LD, Pelletier S, Gingras S, et al. Molecular characterization of LC3-associated phagocytosis reveals distinct roles for Rubicon, NOX2 and autophagy proteins. Nat Cell Biol. 2015;17(7):893–906.
102. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science. 1990;249(4968):505–10.
103. Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods. 2011;8(9):765–70. Pubmed Central PMCID: 3164905.
104. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol. 2011;29(9):816–23.
105. Smith C, Gore A, Yan W, Abalde-Atristain L, Li Z, He C, et al. Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. Cell Stem Cell. 2014;15(1):12–3. Pubmed Central PMCID: 4338993.
106. Iyer V, Shen B, Zhang W, Hodgkins A, Keane T, Huang X, et al. Off-target mutations are rare in Cas9-modified mice. Nat Methods. 2015;12(6):479.
107. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell. 2013;154(6):1370–9. Pubmed Central PMCID: 3961003.
108. Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, et al. Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell. 2014;15(1):27–30. Pubmed Central PMCID: 4082799.
109. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013;153(4):910–8. Pubmed Central PMCID: 3969854.
110. Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol. 2007;25(7):778–85.
111. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol. 2013;31(9):833–8. Pubmed Central PMCID: 3818127.
112. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013;154(6):1380–9. Pubmed Central PMCID: 3856256.
113. Fu Y, Reyon D, Joung JK. Targeted genome editing in human cells using CRISPR/Cas nucleases and truncated guide RNAs. Methods Enzymol. 2014;546:21–45.
114. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat Biotechnol. 2014;32(6):577–82. Pubmed Central PMCID: 4263420.
115. Chapman JR, Taylor MR, Boulton SJ. Playing the end game: DNA double-strand break repair pathway choice. Mol Cell. 2012;47(4):497–510.
116. Chang YF, Imam JS, Wilkinson MF. The nonsense-mediated decay RNA surveillance pathway. Annu Rev Biochem. 2007;76:51–74.

117. Yang H, Wang H, Jaenisch R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. Nat Protoc. 2014;9(8):1956–68.
118. Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res. 2010;38(15), e152. Pubmed Central PMCID: 2926620.
119. Cai CQ, Doyon Y, Ainley WM, Miller JC, Dekelver RC, Moehle EA, et al. Targeted transgene integration in plant cells using designed zinc finger nucleases. Plant Mol Biol. 2009;69(6):699–709.
120. Moehle EA, Rock JM, Lee YL, Jouvenot Y, DeKelver RC, Gregory PD, et al. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci U S A. 2007;104(9):3055–60. Pubmed Central PMCID: 1802009.
121. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011;29(8):731–4. Pubmed Central PMCID: 3152587.
122. Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell. 2010;140(5):678–91. Pubmed Central PMCID: 2885838.
123. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D. Efficient gene targeting in *Drosophila* with zinc-finger nucleases. Genetics. 2006;172(4):2391–403. Pubmed Central PMCID: 1456366.
124. Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, et al. Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci U S A. 2012;109(43):17382–7. Pubmed Central PMCID: 3491456.
125. Ma S, Zhang S, Wang F, Liu Y, Liu Y, Xu H, et al. Highly efficient and specific genome editing in silkworm using custom TALENs. PLoS One. 2012;7(9), e45035. Pubmed Central PMCID: 3445556.
126. Liu PQ, Chan EM, Cost GJ, Zhang L, Wang J, Miller JC, et al. Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. Biotechnol Bioeng. 2010;106(1):97–105.
127. Kim S, Lee HJ, Kim E, Kim JS. Analysis of targeted chromosomal deletions induced by zinc finger nucleases. Cold Spring Harbor Protoc 2010;2010(8):pdb prot5477.
128. Lee HJ, Kim E, Kim JS. Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome Res. 2010;20(1):81–9. Pubmed Central PMCID: 2798833.
129. Petolino JF, Worden A, Curlee K, Connell J, Strange Moynahan TL, Larsen C, et al. Zinc finger nuclease-mediated transgene deletion. Plant Mol Biol. 2010;73(6):617–28.
130. Blasco RB, Karaca E, Ambrogio C, Cheong TC, Karayol E, Minero VG, et al. Simple and rapid in vivo generation of chromosomal rearrangements using CRISPR/Cas9 technology. Cell Rep. 2014;9(4):1219–27.
131. Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han YC, et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature. 2014;516(7531):423–7. Pubmed Central PMCID: 4270925.
132. Ghezraoui H, Piganeau M, Renouf B, Renaud JB, Sallmyr A, Ruis B, et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol Cell. 2014;55(6):829–42. Pubmed Central PMCID: 4398060.
133. Choi PS, Meyerson M. Targeted genomic rearrangements using CRISPR/Cas technology. Nat Commun. 2014;5:3728. Pubmed Central PMCID: 4170920.
134. Torres R, Martin MC, Garcia A, Cigudosa JC, Ramirez JC, Rodriguez-Perales S. Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. Nat Commun. 2014;5:3964.
135. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, et al. Cancer translocations in human cells induced by zinc finger and TALE nucleases. Genome Res. 2013;23(7):1182–93. Pubmed Central PMCID: 3698511.
136. Do TU, Ho B, Shih SJ, Vaughan A. Zinc Finger Nuclease induced DNA double stranded breaks and rearrangements in MLL. Mutat Res. 2012;740(1-2):34–42. Pubmed Central PMCID: 3578303.

137. Van Deursen J, Fornerod M, Van Rees B, Grosveld G. Cre-mediated site-specific translocation between nonhomologous mouse chromosomes. Proc Natl Acad Sci U S A. 1995;92(16):7376–80. Pubmed Central PMCID: 41342.
138. DeKelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijsing SH, et al. Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Res. 2010;20(8):1133–42. Pubmed Central PMCID: 2909576.
139. Luo Y, Rao M, Zou J. Generation of GFP reporter human induced pluripotent stem cells using AAVS1 safe harbor transcription activator-like effector nuclease. Curr Protoc Stem Cell Biol. 2014;29:5A 71–5A 718. Pubmed Central PMCID: 4128243.
140. Tiyaboonchai A, Mac H, Shamsedeen R, Mills JA, Kishore S, French DL, et al. Utilization of the AAVS1 safe harbor locus for hematopoietic specific transgene expression and gene knockdown in human ES cells. Stem Cell Res. 2014;12(3):630–7. Pubmed Central PMCID: 4048956.
141. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol. 2015;33(5):543–8.
142. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol. 2015;33(5):538–42.
143. Singh P, Schimenti JC, Bolcun-Filas E. A mouse geneticist's practical guide to CRISPR applications. Genetics. 2015;199(1):1–15. Pubmed Central PMCID: 4286675.
144. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife. 2014;3, e04766. Pubmed Central PMCID: 4383097.
145. Liu J, Gaj T, Patterson JT, Sirk SJ, Barbas III CF. Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. PLoS One. 2014;9(1), e85755. Pubmed Central PMCID: 3896395.
146. Ru R, Yao Y, Yu S, Yin B, Xu W, Zhao S, et al. Targeted genome engineering in human induced pluripotent stem cells by penetrating TALENs. Cell Regen. 2013;2(1):5. Pubmed Central PMCID: 4230761.
147. Chen Z, Jaafar L, Agyekum DG, Xiao H, Wade MF, Kumaran RI, et al. Receptor-mediated delivery of engineered nucleases for genome modification. Nucleic Acids Res. 2013;41(19), e182. Pubmed Central PMCID: 3799454.
148. Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome Res. 2014;24(6):1020–7. Pubmed Central PMCID: 4032848.
149. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas III CF. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods. 2012;9(8):805–7. Pubmed Central PMCID: 3424280.
150. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 2014;24(6):1012–9. Pubmed Central PMCID: 4032847.
151. Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotechnol. 2015;33(1):73–80. Pubmed Central PMCID: 4289409.
152. Sahin U, Kariko K, Tureci O. mRNA-based therapeutics—developing a new class of drugs. Nat Rev Drug Discov. 2014;13(10):759–80.
153. Genovese P, Schiroli G, Escobar G, Di Tomaso T, Firrito C, Calabria A, et al. Targeted genome editing in human repopulating haematopoietic stem cells. Nature. 2014;510(7504):235–40. Pubmed Central PMCID: 4082311.
154. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol. 2007;25(11):1298–306.

155. Provasi E, Genovese P, Lombardo A, Magnani Z, Liu PQ, Reik A, et al. Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. Nat Med. 2012;18(5):807–15.
156. Holkers M, Maggio I, Liu J, Janssen JM, Miselli F, Mussolino C, et al. Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic Acids Res. 2013;41(5), e63. Pubmed Central PMCID: 3597656.
157. Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, Mali P, et al. Optimization of scarless human stem cell genome editing. Nucleic Acids Res. 2013;41(19):9049–61. Pubmed Central PMCID: 3799423.
158. Mock U, Riecken K, Berdien B, Qasim W, Chan E, Cathomen T, et al. Novel lentiviral vectors with mutated reverse transcriptase for mRNA delivery of TALE nucleases. Sci Rep. 2014;4:6409. Pubmed Central PMCID: 4166709.
159. Lei Y, Lee CL, Joo KI, Zarzar J, Liu Y, Dai B, et al. Gene editing of human embryonic stem cells via an engineered baculoviral vector carrying zinc-finger nucleases. Mol Ther. 2011;19(5):942–50. Pubmed Central PMCID: 3098635.
160. Phang RZ, Tay FC, Goh SL, Lau CH, Zhu H, Tan WK, et al. Zinc finger nuclease-expressing baculoviral vectors mediate targeted genome integration of reprogramming factor genes to facilitate the generation of human induced pluripotent stem cells. Stem Cells Transl Med. 2013;2(12):935–45. Pubmed Central PMCID: 3841088.
161. Tay FC, Tan WK, Goh SL, Ramachandra CJ, Lau CH, Zhu H, et al. Targeted transgene insertion into the AAVS1 locus driven by baculoviral vector-mediated zinc finger nuclease expression in human-induced pluripotent stem cells. J Gene Med. 2013;15(10):384–95.
162. Zhu H, Lau CH, Goh SL, Liang Q, Chen C, Du S, et al. Baculoviral transduction facilitates TALEN-mediated targeted transgene integration and Cre/LoxP cassette exchange in human-induced pluripotent stem cells. Nucleic Acids Res. 2013;41(19), e180. Pubmed Central PMCID: 3799456.
163. Lau CH, Zhu H, Tay JC, Li Z, Tay FC, Chen C, et al. Genetic rearrangements of variable di-residue (RVD)-containing repeat arrays in a baculoviral TALEN system. Mol Ther Methods Clin Dev. 2014;1:14050. Pubmed Central PMCID: 4362386.
164. Wirth T, Parker N, Yla-Herttuala S. History of gene therapy. Gene. 2013;525(2):162–9.
165. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med. 2014;370(10):901–10. Pubmed Central PMCID: 4084652.
166. Coluccio A, Miselli F, Lombardo A, Marconi A, Malagoli Tagliazucchi G, Goncalves MA, et al. Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. Mol Ther. 2013;21(9):1695–704. Pubmed Central PMCID: 3776632.
167. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol. 2008;26(7):808–16. Pubmed Central PMCID: 3422503.
168. Holkers M, Maggio I, Henriques SF, Janssen JM, Cathomen T, Goncalves MA. Adenoviral vector DNA for accurate genome editing with engineered nucleases. Nat Methods. 2014;11(10):1051–7.
169. Maggio I, Holkers M, Liu J, Janssen JM, Chen X, Goncalves MA. Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. Sci Rep. 2014;4:5105. Pubmed Central PMCID: 4037712.
170. Holkers M, Cathomen T, Goncalves MA. Construction and characterization of adenoviral vectors for the delivery of TALENs into human cells. Methods. 2014;69(2):179–87.
171. Kotterman MA, Schaffer DV. Engineering adeno-associated viruses for clinical gene therapy. Nat Rev Genet. 2014;15(7):445–51. Pubmed Central PMCID: 4393649.
172. Ellis BL, Hirsch ML, Porter SN, Samulski RJ, Porteus MH. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. Gene Ther. 2013;20(1):35–42.

173. Asuri P, Bartel MA, Vazin T, Jang JH, Wong TB, Schaffer DV. Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. Mol Ther. 2012;20(2):329–38. Pubmed Central PMCID: 3277219.
174. Handel EM, Gellhaus K, Khan K, Bednarski C, Cornu TI, Muller-Lerch F, et al. Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors. Hum Gene Ther. 2012;23(3):321–9. Pubmed Central PMCID: 3300077.
175. Howes R, Schofield C. Genome engineering using Adeno-Associated Virus (AAV). Methods Mol Biol. 2015;1239:75–103.
176. Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell. 2014;159(2):440–55. Pubmed Central PMCID: 4265475.
177. Senis E, Fatouros C, Grosse S, Wiedtke E, Niopek D, Mueller AK, et al. CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. Biotechnol J. 2014;9(11):1402–12.
178. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 2013;41(7):4336–43. Pubmed Central PMCID: 3627607.
179. Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, et al. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res. 2011;39(14):6315–25. Pubmed Central PMCID: 3152341.
180. Morton J, Davis MW, Jorgensen EM, Carroll D. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. Proc Natl Acad Sci U S A. 2006;103(44):16370–5. Pubmed Central PMCID: 1637589.
181. Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM, Calarco JA. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nat Methods. 2013;10(8):741–3. Pubmed Central PMCID: 3822328.
182. Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, et al. Targeted genome editing across species using ZFNs and TALENs. Science. 2011;333(6040):307. Pubmed Central PMCID: 3489282.
183. Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, et al. Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. J Genet Genomics. 2012; 39(5):209–15.
184. Bibikova M, Golic M, Golic KG, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. Genetics. 2002;161(3):1169–75. Pubmed Central PMCID: 1462166.
185. Bassett AR, Tibbit C, Ponting CP, Liu JL. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. Cell Rep. 2013;4(1):220–8. Pubmed Central PMCID: 3714591.
186. Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, et al. Heritable and precise zebrafish genome editing using a CRISPR-Cas system. PLoS One. 2013;8(7), e68708. Pubmed Central PMCID: 3706373.
187. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. 2013;31(3):227–9. Pubmed Central PMCID: 3686313.
188. Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, et al. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol. 2011;29(8):697–8. Pubmed Central PMCID: 3154023.
189. Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, et al. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol. 2008;26(6):702–8. Pubmed Central PMCID: 2674762.
190. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat Biotechnol. 2008;26(6):695–701. Pubmed Central PMCID: 2502069.

191. Young JJ, Cherone JM, Doyon Y, Ankoudinova I, Faraji FM, Lee AH, et al. Efficient targeted gene disruption in the soma and germ line of the frog *Xenopus tropicalis* using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A. 2011;108(17):7052–7. Pubmed Central PMCID: 3084115.
192. Lei Y, Guo X, Liu Y, Cao Y, Deng Y, Chen X, et al. Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). Proc Natl Acad Sci U S A. 2012;109(43):17484–9. Pubmed Central PMCID: 3491516.
193. Blitz IL, Biesinger J, Xie X, Cho KW. Biallelic genome modification in F(0) *Xenopus tropicalis* embryos using the CRISPR/Cas system. Genesis. 2013;51(12):827–34. Pubmed Central PMCID: 4039559.
194. Nakayama T, Fish MB, Fisher M, Oomen-Hajagos J, Thomsen GH, Grainger RM. Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. Genesis. 2013;51(12):835–43. Pubmed Central PMCID: 3947545.
195. Carbery ID, Ji D, Harrington A, Brown V, Weinstein EJ, Liaw L, et al. Targeted genome modification in mice using zinc-finger nucleases. Genetics. 2010;186(2):451–9. Pubmed Central PMCID: 2954478.
196. Meyer M, de Angelis MH, Wurst W, Kuhn R. Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proc Natl Acad Sci U S A. 2010;107(34):15022–6. Pubmed Central PMCID: 2930558.
197. Qiu Z, Liu M, Chen Z, Shao Y, Pan H, Wei G, et al. High-efficiency and heritable gene targeting in mouse by transcription activator-like effector nucleases. Nucleic Acids Res. 2013;41(11), e120. Pubmed Central PMCID: 3675477.
198. Wefers B, Panda SK, Ortiz O, Brandl C, Hensler S, Hansen J, et al. Generation of targeted mouse mutants by embryo microinjection of TALEN mRNA. Nat Protoc. 2013;8(12):2355–79.
199. Wefers B, Meyer M, Ortiz O, Hrabe de Angelis M, Hansen J, Wurst W, et al. Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. Proc Natl Acad Sci U S A. 2013;110(10):3782–7. Pubmed Central PMCID: 3593923.
200. Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. Science. 2009;325(5939):433. Pubmed Central PMCID: 2831805.
201. Tesson L, Usal C, Menoret S, Leung E, Niles BJ, Remy S, et al. Knockout rats generated by embryo microinjection of TALENs. Nat Biotechnol. 2011;29(8):695–6.
202. Li W, Teng F, Li T, Zhou Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. Nat Biotechnol. 2013;31(8):684–6.
203. Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M, et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. Nat Biotechnol. 2013;31(8):681–3.
204. Flisikowska T, Thorey IS, Offner S, Ros F, Lifke V, Zeitler B, et al. Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. PLoS One. 2011;6(6), e21045. Pubmed Central PMCID: 3113902.
205. Song J, Zhong J, Guo X, Chen Y, Zou Q, Huang J, et al. Generation of RAG 1- and 2-deficient rabbits by embryo microinjection of TALENs. Cell Res. 2013;23(8):1059–62. Pubmed Central PMCID: 3731565.
206. Yan Q, Zhang Q, Yang H, Zou Q, Tang C, Fan N, et al. Generation of multi-gene knockout rabbits using the Cas9/gRNA system. Cell Regen. 2014;3(1):12. Pubmed Central PMCID: 4230364.
207. Honda A, Hirose M, Sankai T, Yasmin L, Yuzawa K, Honsho K, et al. Single-step generation of rabbits carrying a targeted allele of the tyrosinase gene using CRISPR/Cas9. Exp Anim. 2015;64(1):31–7. Pubmed Central PMCID: 4329513.
208. Chen Y, Zheng Y, Kang Y, Yang W, Niu Y, Guo X, et al. Functional disruption of the dystrophin gene in rhesus monkey using CRISPR/Cas9. Hum Mol Genet. 2015;24(13):3764–74.
209. Wan H, Feng C, Teng F, Yang S, Hu B, Niu Y, et al. One-step generation of p53 gene biallelic mutant *Cynomolgus* monkey via the CRISPR/Cas system. Cell Res. 2015;25(2):258–61.
210. Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, et al. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell. 2014;156(4):836–43.

211. Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, et al. High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. Proc Natl Acad Sci U S A. 2010;107(26):12028–33. Pubmed Central PMCID: 2900673.
212. Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, et al. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature. 2009;459(7245):437–41.
213. Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, et al. High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature. 2009; 459(7245):442–5. Pubmed Central PMCID: 2743854.
214. Li T, Liu B, Spalding MH, Weeks DP, Yang B. High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol. 2012;30(5):390–2.
215. Jia H, Wang N. Targeted genome editing of sweet orange using Cas9/sgRNA. PLoS One. 2014;9(4), e93806. Pubmed Central PMCID: 3977896.
216. Nodvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. PLoS One. 2015;10(7), e0133085. Pubmed Central PMCID: 4503723.
217. Arazoe T, Miyoshi K, Yamato T, Ogawa T, Ohsato S, Arie T, et al. Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng. 2015;112:2543–9.
218. Wu H, Wang Y, Zhang Y, Yang M, Lv J, Liu J, et al. TALE nickase-mediated SP110 knockin endows cattle with increased resistance to tuberculosis. Proc Natl Acad Sci U S A. 2015;112(13):E1530–9. Pubmed Central PMCID: 4386332.
219. Tan W, Carlson DF, Lancto CA, Garbe JR, Webster DA, Hackett PB, et al. Efficient nonmeiotic allele introgression in livestock using custom endonucleases. Proc Natl Acad Sci U S A. 2013;110(41):16526–31. Pubmed Central PMCID: 3799378.
220. Proudfoot C, Carlson DF, Huddart R, Long CR, Pryor JH, King TJ, et al. Genome edited sheep and cattle. Transgenic Res. 2015;24(1):147–53. Pubmed Central PMCID: 4274373.
221. Lillico SG, Proudfoot C, Carlson DF, Stverakova D, Neil C, Blain C, et al. Live pigs produced from genome edited zygotes. Sci Rep. 2013;3:2847.
222. Estrada JL, Martens G, Li P, Adams A, Newell KA, Ford ML, et al. Evaluation of human and non-human primate antibody binding to pig cells lacking GGTA1/CMAH/beta4GalNT2 genes. Xenotransplantation. 2015;22(3):194–202. Pubmed Central PMCID: 4464961.
223. Li P, Estrada JL, Burlak C, Montgomery J, Butler JR, Santos RM, et al. Efficient generation of genetically distinct pigs in a single pregnancy using multiplexed single-guide RNA and carbohydrate selection. Xenotransplantation. 2015;22(1):20–31.
224. Kang JT, Kwon DK, Park AR, Lee EJ, Yun YJ, Ji DY, et al. Production of GGTA1 targeted pigs using TALEN-mediated genome editing technology. J Vet Sci. 2016;17(1):89–96. PMCID: 4808648.
225. Xin J, Yang H, Fan N, Zhao B, Ouyang Z, Liu Z, et al. Highly efficient generation of GGTA1 biallelic knockout inbred mini-pigs with TALENs. PLoS One. 2013;8(12), e84250. Pubmed Central PMCID: 3866186.
226. Prather RS, Lorson M, Ross JW, Whyte JJ, Walters E. Genetically engineered pig models for human diseases. Annu Rev Anim Biosci. 2013;1:203–19. Pubmed Central PMCID: 4460601.
227. Liu X, Wang Y, Tian Y, Yu Y, Gao M, Hu G, et al. Generation of mastitis resistance in cows by targeting human lysozyme gene to beta-casein locus using zinc-finger nucleases. Proc Biol Sci. 2014;281(1780):20133368. Pubmed Central PMCID: 4027401.
228. Crispo M, Mulet AP, Tesson L, Barrera N, Cuadro F, Dos Santos-Neto PC, et al. Efficient generation of myostatin knock-out sheep using CRISPR/Cas9 technology and microinjection into zygotes. PLoS One. 2015;10(8), e0136690. Pubmed Central PMCID: 4549068.
229. Sebastiano V, Maeder ML, Angstman JF, Haddad B, Khayter C, Yeo DT, et al. In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. Stem Cells. 2011;29(11):1717–26. Pubmed Central PMCID: 3285772.

230. Finotti A, Breda L, Lederer CW, Bianchi N, Zuccato C, Kleanthous M, et al. Recent trends in the gene therapy of beta-thalassemia. J Blood Med. 2015;6:69–85. Pubmed Central PMCID: 4342371.
231. Huang X, Wang Y, Yan W, Smith C, Ye Z, Wang J, et al. Production of gene-corrected adult beta globin protein in human erythrocytes differentiated from patient iPSCs after genome editing of the sickle point mutation. Stem Cells. 2015;33(5):1470–9.
232. Li H, Haurigot V, Doyon Y, Li T, Wong SY, Bhagwat AS, et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. Nature. 2011;475(7355):217–21. Pubmed Central PMCID: 3152293.
233. Anguela XM, Sharma R, Doyon Y, Miller JC, Li H, Haurigot V, et al. Robust ZFN-mediated genome editing in adult hemophilic mice. Blood. 2013;122(19):3283–7. Pubmed Central PMCID: 3821724.
234. Park CY, Kim DH, Son JS, Sung JJ, Lee J, Bae S, et al. Functional correction of large factor VIII gene chromosomal inversions in hemophilia A patient-derived iPSCs using CRISPR-Cas9. Cell Stem Cell. 2015;17(2):213–20.
235. Park CY, Kim J, Kweon J, Son JS, Lee JS, Yoo JE, et al. Targeted inversion and reversion of the blood coagulation factor 8 gene in human iPS cells using TALENs. Proc Natl Acad Sci U S A. 2014;111(25):9253–8. Pubmed Central PMCID: 4078797.
236. Smith C, Abalde-Atristain L, He C, Brodsky BR, Braunstein EM, Chaudhari P, et al. Efficient and allele-specific genome editing of disease loci in human iPSCs. Mol Ther. 2015;23(3):570–7. Pubmed Central PMCID: 4351458.
237. Ousterout DG, Kabadi AM, Thakore PI, Majoros WH, Reddy TE, Gersbach CA. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. Nat Commun. 2015;6:6244. Pubmed Central PMCID: 4335351.
238. Li HL, Fujimoto N, Sasakawa N, Shirai S, Ohkame T, Sakuma T, et al. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. Stem Cell Rep. 2015;4(1):143–54. Pubmed Central PMCID: 4297888.
239. Toscano MG, Anderson P, Munoz P, Lucena G, Cobo M, Benabdellah K, et al. Use of zinc-finger nucleases to knock out the WAS gene in K562 cells: a human cellular model for Wiskott-Aldrich syndrome. Dis Model Mech. 2013;6(2):544–54. Pubmed Central PMCID: 3597037.
240. Firth AL, Menon T, Parker GS, Qualls SJ, Lewis BM, Ke E, et al. Functional gene correction for cystic fibrosis in lung epithelial cells generated from patient iPSCs. Cell Rep. 2015;12(9):1385–90. Pubmed Central PMCID: 4559351.
241. Ding Q, Strong A, Patel KM, Ng SL, Gosis BS, Regan SN, et al. Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. Circ Res. 2014;115(5):488–92. Pubmed Central PMCID: 4134749.
242. Long C, McAnally JR, Shelton JM, Mireault AA, Bassel-Duby R, Olson EN. Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science. 2014;345(6201):1184–8. Pubmed Central PMCID: 4398027.
243. Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, et al. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol. 2010;28(8):839–47. Pubmed Central PMCID: 3080757.
244. Hu W, Kaminski R, Yang F, Zhang Y, Cosentino L, Li F, et al. RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. Proc Natl Acad Sci U S A. 2014;111(31):11461–6. Pubmed Central PMCID: 4128125.
245. Yuen KS, Chan CP, Wong NH, Ho CH, Ho TH, Lei T, et al. CRISPR/Cas9-mediated genome editing of Epstein-Barr virus in human cells. J Gen Virol. 2015;96(Pt 3):626–36.
246. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. Nature. 2015;520(7546):186–91. Pubmed Central PMCID: 4393360.

CRISPR/Cas9 Approaches to Investigate the Noncoding Genome

Davide Seruggia and Lluis Montoliu

Abstract Studies of gene function in mice have been supported during the past decade by a nearly exhaustive collection of mutants, systematically obtained by homologous recombination in murine ES cells. Unfortunately, the study of the noncoding fraction of the genome did not benefit from the same valuable resources. Nevertheless, increasing evidence of the relevance of this fraction of the vertebrate genome has been accumulated in the past years. Comprehensive maps of histone modifications, methylation patterns, and DNA-binding protein occupancies have been made available to predict key regulatory elements through the work of various international collaborative consortia, such as ENCODE. Comparing these maps with data from genome-wide association studies (GWAS) suggested that variants in noncoding sequence elements might be involved in several traits and disease conditions. Therefore, there is an urgent need for accurate functional tests and genetic modelling of noncoding elements. In this chapter, we propose a number of strategies to test hypothesis regarding noncoding DNA elements, by taking advantage of the most recent genome editing techniques, namely, CRISPR/Cas9 approaches.

Keywords Enhancer • Insulator • ChIP-seq • Transcription factor • Chromatin • Epigenetics

D. Seruggia
Department of Molecular and Cellular Biology, National Centre for Biotechnology (CNB-CSIC), Campus Cantoblanco, Darwin 3, 28049 Madrid, Spain

CIBERER-ISCIII, Madrid, Spain

Department of Hematology/Oncology, Boston Children's Hospital, Karp Family Research Laboratories 7, 1 Blackfan Circle, Boston, MA 02115, USA

L. Montoliu (✉)
Department of Molecular and Cellular Biology, National Centre for Biotechnology (CNB-CSIC), Campus Cantoblanco, Darwin 3, 28049 Madrid, Spain

CIBERER-ISCIII, Madrid, Spain
e-mail: montoliu@cnb.csic.es

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_2

Introduction

Protein-coding genes account for just a small fraction of the vertebrate genomes, estimated to be about 2 %. The remainder, 98 % of the genome, is composed of repetitive DNA elements and, most importantly, of a large variety of noncoding sequences that share a considerable degree of conservation across species [1]. Such conserved, noncoding elements are required for accurate regulation of gene expression and are involved in determining cell type identity and function. Active and inactive regulatory sequences are associated with specific biochemical marks. DNA methylation, histone modifications, and protein occupancy are predictive for the function and the state of a given element [2]. Active enhancer sequences are typically marked by mono-methylation of lysine 4 of histone H3 (H3K4me1), by acetylation of lysine 27 (H3K27ac), and localize in open chromatin regions that show hypersensitivity to DNase I digestion (Fig. 1). In addition, the distribution of two enhancer-binding proteins, CHD7 and P300 [3], can be used with confidence to identify enhancers. Binding profiles of transcription factors in gene neighbourhoods allow inferring the underlying networks of gene regulation [4]. Interestingly, the distribution of chromatin marks and transcription factors that decorates the noncoding fraction of the genome varies between different cell types and changes dynamically during development and differentiation [5]. A direct correlation between transcription profile and chromatin signatures has been described in many cell types or tissues, including embryonic stem (ES) cells [6], further triggering the interest for noncoding, regulatory elements in the fields of development and stem cell biology. Several sequence variants in noncoding elements have already been found associated with human traits as well as with disease conditions [7–11]. For example, genome-wide association studies (GWAS) highlighted that a common trait such as eye colour is strongly associated with a DNA polymorphism lying 21 kb upstream of the pigmentation-related OCA2 gene. Molecular analyses indicated that a particular single-nucleotide polymorphism (SNP) is located within a OCA2 enhancer and, interestingly, the rs12913832 C-allele is associated with decreased OCA2 expression, reduced transcription factor recruitment, and chromatin looping [11]. Hair colour also results, in part, from variants at noncoding sequences [12]. Thus, increasing evidence denotes the functional role of noncoding variants in both human traits and disease. Therefore, there is a pressing need of adequate modelling of noncoding variants and mutations. This chapter aims to provide an overview of the experimental approaches that can be used to study the role of noncoding DNA elements and to obtain their inactivation in model systems using the CRISPR/Cas9 system.

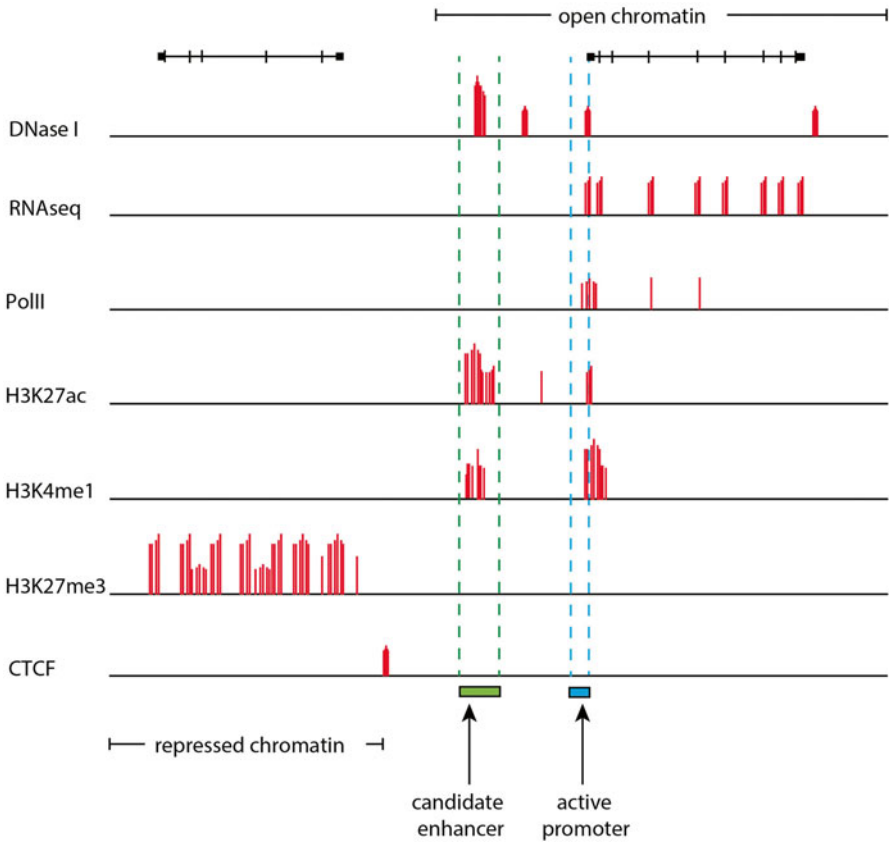

Fig. 1 Chromatin marks guide the identification of active regulatory elements. ENCODE and Roadmap Epigenomics datasets can be browsed to identify putative regulatory elements. DNaseI-seq tracks highlight open chromatin regions. RNA-seq tracks can be used to identify actively transcribing genes. Similarly, RNA polymerase II (PolII) ChIP-seq tracks marks transcription. H3K27ac is a typical mark of active enhancers, whereas H3K4me1 can be used to identify active promoters. In contrast, H3K27me3 marks repressed chromatin. CTCF protein is associated with chromatin boundaries and marks the transition between open and closed chromatin

Genetic Manipulations of Noncoding Sequences in the Prenuclease Era

Loss-of-function alleles of nearly all protein-coding genes have been obtained by homologous recombination in mouse ES cells [13]. On paper, the same experimental approach could be applied to inactivate noncoding elements. However, only a handful of noncoding elements have been inactivated using this strategy in ES cells [14–16]. In fact, the noncoding fraction of the genome is particularly enriched in repetitive elements, a characteristic that poses difficulties and challenges the design of optimal targeting vectors, because the presence of non-unique DNA sequences

within the homology arms is obviously detrimental for targeting efficiency. Hence, most targeting events would be expected to occur elsewhere, outside the desired noncoding DNA element, but at genomic locations rich in similar repetitive sequences.

Furthermore, a recent study highlighted a bias associated with gene targeting in ES cells. In fact, by comparing the genome of 129/Sv and C57BL/6J, the mouse strains that are most commonly used as source of ES cells and as recipient embryos, respectively, more than 1000 passenger mutations in the vicinity of coding genes have been identified, which can confound the interpretation of the associated mutant phenotypes [17].

An alternative strategy that has been largely explored in previous years involves the use of large, genomic-type, transgenes [18]. Instead of targeting regulatory elements at their endogenous locus, several laboratories reproduced inactivating mutations within large genomic constructs, such as those included with artificial chromosome type of transgenes (i.e., bacterial artificial chromosomes, BACs; or yeast artificial chromosomes, YACs). There, the desired mutation is built in the context of a large genomic-type transgene. Using bacteria or yeast, the efficiency of homologous recombination is much higher and the handling of large numbers of clones is easier and also less expensive [19]. In addition, the recombined constructs are delivered to cells or model organisms, and the effect of the mutation in noncoding DNA elements is often read through the activity of a reporter gene included in the transgene [20]. Alternatively, the modified BAC or YAC can be introduced in mice where the endogenous gene was previously inactivated [21]. Nevertheless, this approach suffers from a number of drawbacks. First, large constructs are not easy to manipulate and not all laboratories succeeded in establishing the required protocols. Second, modelling a mutation in a large transgene would in turn introduce several other non-isogenic variants into the model. Finally, copy number, site of integration, and transgene integrity may severely affect the phenotype, leading to complex phenotypes resulting from a mixture of variables, where variegated expression of the transgenes can be also a confounding factor [22, 23].

The CRISPR/Cas9 system, with its high efficiency and flexibility, seems to be the ideal candidate to fill the existing gap in the genetic modelling and functional assessment of noncoding sequences in vitro and in nearly all model organisms [24, 25]. By using this system, the limitations that we described associated with gene targeting in ES cells and large transgenes can be easily overcome. In particular, with the use of CRISPR/Cas9 approaches the requirement for sequence homology can be reduced to just 20 base pairs, relatively easy to find even within stretches of repetitive DNA sequences. And, most importantly, the modification occurs precisely at the endogenous locus, hence avoiding any chromosomal position effect [23].

In this chapter we illustrate three distinct strategies for the genetic perturbation of noncoding elements using CRISPR/Cas9 approaches. First, we introduce the inactivation by chromosomal deletion (Fig. 2a). Next, we discuss a novel technique known as *epigenome editing* (Fig. 2b). Finally, we discuss strategies to target particular non-protein-coding genes, such as miRNA and lincRNAs.

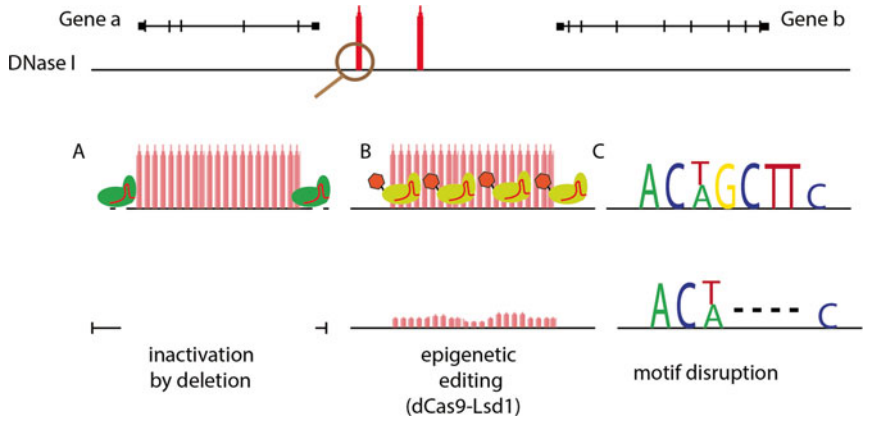

Fig. 2 Different CRISPR/Cas9 approaches. (**a**) Inactivation by deletion: two single guide (sg)RNAs are designed to flank a chromatin region of interest. When simultaneous double-stranded breaks are produced, the intervening DNA is deleted. (**b**) Epigenetic editing: sgRNAs are designed across the target sequence and delivered with a dCas9 protein fused with the desired chromatin remodelling catalytic domain. For example, the Lsd1 catalytic domain is able to inactivate active enhancers. (**c**) Motif disruption. When the DNA-binding specificities of a transcription factor are well characterized and defined, a single sgRNA can be devised to target specifically the nucleotides that constitute a specific DNA-binding site. Upon NHEJ DNA repair, *indel* mutations will be introduced at the DNA-binding motif, potentially abolishing the binding of the cognate transcription factor

Editing by Chromosomal Deletion

Noncoding regulatory elements are contained within arrays of transcription factor-binding sites. Their mechanism of action often relies on the cooperative occupancy of multiple DNA-binding proteins. Binding of a specific transcription factor to a DNA sequence can be anticipated by the presence of a particular motif or consensus sequence. However, DNA-binding motifs are not always completely predictive because many consensus-binding sequences are poorly defined. Position weight matrices (PWM)-based prediction of DNA-binding consensus highlighted that the majority of transcription factors tolerate multiple point mutations at their target sequence [26]. Hence, different from protein-coding genes that can be disrupted introducing frameshift mutations at their sequence, point mutations might be insufficient to fully inactivate complex regulatory elements. One thoughtful approach would be to remove the noncoding DNA element from the genome, within a larger DNA sequence, including multiple predicted DNA-binding sites, that is hypothetically linked with a function (typically, a 1- to 2-kb large sequence). Next, in a second round of experiments, smaller overlapping deletions can be produced to fine-map core elements or to dissect the differential relevance associated with each of the individual DNA sequences. The nonhomologous end-joining DNA repair route (NHEJ), triggered by the CRISPR/Cas9 approach, can be exploited to obtain such alleles, promoting religation of two distal DNA ends generated by adjacent and

simultaneous double-strand breaks (DSB); this would in turn lead to the loss of the intervening DNA sequences existing between the two DSBs (Fig. 2a). This method represents a simple and effective approach to obtain the targeted inactivation of noncoding elements, in their endogenous genomic context, by designing two targeted nucleases (i.e., in a CRISPR/Cas9 approach: two single guide RNAs, sgRNA, to drive the DSB caused by Cas9). Through surgical removal of the putative regulatory element, alteration in gene expression level of nearby and distal genes can be measured [27], as well as alterations in chromatin marks at the surrounding sequences.

In our laboratory, we have used a CRISPR/Cas9 deletion approach to test the in vivo relevance of a regulatory element found upstream from the mouse *Tyr* gene (encoding tyrosinase, the first and fundamental enzyme in the biosynthetic pathway of melanin) that we had previously investigated using YAC-based and standard transgenes [21, 28, 29]. The *Tyr* 5′-upstream region contains multiple DNA-binding motifs and it is marked by EP300, a DNA-binding protein that often decorates active enhancers, and by H3K4me1 [23]. Through delivery of two sgRNAs flanking this element and the Cas9 mRNA to mouse fertilized eggs, we have generated several deletion alleles with high frequency [23]. When homozygous deletions are produced, a similar loss of coat-colour pigmentation is observed in several independent lines, indicating that such an element is indeed required to achieve wild-type *Tyr* expression. As stated before, double-CRISPR/Cas9 deletions are generated through the error-prone NHEJ DNA repair route. Thus, each allele we obtained carries a typical and unique scar at the DNA sealing point. Another consequence of error-prone DNA repair is the production of partial and larger deletions, and even inversions, probably the result of rearrangements during DNA repair, favoured by the number of repetitive DNA elements that flank the target sequences. Such additional alleles can be used for genetic mapping. In fact, by comparing the phenotype of distinct alleles, we have assembled a genetic map of this mouse *Tyr* 5′-enhancer/boundary element [23].

A previous study in cultured cells highlighted that the efficiency of induced chromosomal deletions decreases with the size of the desired deletion [30]. Nevertheless, deletions in the megabase order are indeed possible [31], opening the possibility of precise modelling large chromosomal deletions, structural variants, or copy number variation (CNV) that are often found in families associated with genetic conditions or diseases.

One striking example has been reported by Lupiáñez and colleagues upon modelling mutations at genomic boundaries. Genomic boundaries, or insulators, are regulatory elements involved in the regulation of multiple genes domains. Typically, boundaries are located in between two different but adjacent topological expression domains [32–34]. Genomic boundaries physically separate chromatin territories containing genes with distinct expression patterns and restrict the activity of enhancer clusters to the cognate gene set [34]. Studying such sequences in an ectopic manner, using DNA constructs linked to reporter genes that mimic the natural conformation, might reveal the insulating activity of these boundary elements but does not reveal the true function of such elements in the endogenous context [35]. In contrast, this can be achieved by inactivating them at the endogenous site by

CRISPR/Cas9 deletion [23]. Interestingly, imposing such genomic alterations in mice reproduced the phenotype observed in humans carrying similar chromosomal aberrations [36].

The CRISPR/Cas9-deletion approach can also be applied to the genetic dissection of super- or stretch-enhancers [37]. In fact, sgRNAs can be easily designed to flank the full super-enhancer as well as individual regions associated with discrete chromatin immunoprecipitation resolved by DNA sequencing (ChIP-seq) peaks. With this approach, the contribution of each module of the super-enhancer can be elucidated [38]. For some specific cases, inactivation of noncoding elements can also be achieved using one sgRNA, by disrupting a previously characterized DNA-binding motif (Fig. 2c); this is the case of well-characterized transcription factor-binding motifs, such as GATA-2. Recently, Bresnick and colleagues described the relevance of a number of GATA-2-binding sites in hematopoiesis [39]. By targeting a CRISPR/Cas9 approach to the consensus DNA sequence, *indels* associated with the DNA repair scar will mutate some of the nucleotides constituting such a motif. As a caveat, one must always take into account that several transcription factor-binding sites can still be recognized by the corresponding nuclear factors, even though their sequences might differ significantly from the observed consensus. Because the frequency of CRISPR/Cas9 target sites in the mammalian genome approaches one every eight nucleotides, tiled sgRNAs can also be generated to interrogate a regulatory element under saturating conditions. Using this approach, Canver and colleagues identified vulnerabilities within an erythroid-specific BCL11A enhancer that can be used for therapeutical fetal globin re-induction in the context of sickle cell disease and β-thalassemias [40].

Recently, the mutational signatures of several types of malignancies have been described. Interestingly, mutations at noncoding sequences have been identified in a significant number of cancer patients [41–43]. In contrast to mutations at protein-coding genes, whose impact can be predicted by taking advantage of already existing animal and cellular models, the effect of a mutation outside the coding sequence can be difficult to anticipate. Hence, CRISPR/Cas9 mutagenesis represents a valid experimental approach to rapidly identify the effect of those noncoding mutations. For example, mutations at a noncoding sequence were found accumulated in a number of chronic lymphocytic leukemia (CLL) patients. In fact, a hypermutated region was found to display enhancer-like features such as H3K4 and H3K27ac enrichment. By CRISPR/Cas9-mediated chromosomal deletion of this putative enhancer, Puente and colleagues proved that loss of that enhancer resulted in a 40 % decrease in *Pax5*, a gene located 330 kb upstream [44].

Epigenome Editing

Epigenetic modifications are defined as chemical modifications of either DNA or histone proteins that affect chromatin structure and accessibility to DNA-binding proteins. These modifications are dynamically remodelled during development, differentiation, and aging. In fact, specific classes of enzymes exist that edit back and

forth such biochemical marks, including DNA and histone methylases, demethylases, histone acetyl transferases (HAT), and deacetylases (HDAC) [45]. Each of these enzymes can be targeted by specific drugs to inhibit their activities. In the past years, these inhibitors were used to interfere with chromatin-remodelling enzymes. Furthermore, the use of these molecules was also proposed and has already been explored as potential chemotherapy agents [46–48]. Unfortunately, these inhibitors act globally, nonspecifically. Advances in genome editing and synthetic biology led to the ability of building designer DNA-binding proteins over distinct targeting nuclease platforms. Artificial DNA-binding scaffolds can be coupled with a variety of catalytic domains, including those from the chromatin remodelling proteins listed here. This method allows imposing locally, and in a targeted manner, specific chromatin marks in the absence of any genetic, irreversible manipulation (Fig. 2b). For example, a TALE (transcription activators-like effectors) DNA-binding array was coupled with Tet1, an enzyme that promotes DNA demethylation. Targeting such TALE-Tet1 fusion to methylated DNA targets resulted in the loss of local DNA methylation. For example, by targeting the *RHOXF2* promoter in Hela and HEK 293 cells, this resulted in a 50- to 1000-fold gene activation [49]. Recently, the CRISPR/Cas9 system has been also adapted for applications in epigenome editing, by using a nuclease-dead Cas9 variant (dCas9). By fusing dCas9 with a catalytic domain of LSD1, a lysine-specific demethylase produced a tool to target active enhancers. By programming dCas9-LSD1 with enhancer-specific gRNAs, Kearns and colleagues could induce downregulation of target genes [50]. By targeting enhancers relevant for the expression of pluripotency factors, these authors induced morphological changes in murine ES cells. Interestingly, these changes depend on the depletion of H3K27ac from key enhancers.

It is also possible to turn the switch in the opposite direction. By coupling dCas9 with EP300, a histone acetyltransferase, a strong activation of target genes can be achieved by using enhancer- and promoter-specific guide RNAs [51].

This approach is not associated with irreversible DNA sequence alteration, a feature that is highly desirable for the study of chromatin dynamics. In fact, only the epigenetic features, and not the DNA sequences themselves, are altered. This factor allows performing a sequence-based assay, such as chromosome conformation capture, in the epigenetic-edited cells, a possibility that is obviously lost if a deletion occurs at the targeted sequence in the genome. In addition, reversibility and noninheritability traits are highly desirable features for therapeutical applications. Therefore, epigenome editing is one of the most promising applications derived from the use of the CRISPR/Cas9 tools [52].

Targeting Noncoding RNAs

Recent evidence supported the role of noncoding RNAs, including micro-RNAs (miRNAs) and long intergenic noncoding RNAs (lincRNAs) in regulating gene expression at multiple levels [53]. Catalogues of these transcripts account for more

than 56,000 human lncRNAs [54] and a similar number of small miRNAs. Thus, appropriate genetic models are required to define the role of this emerging class of regulatory RNA molecules. Dissimilar to protein-coding genes, these transcripts cannot be inactivated by triggering frameshift mutations at their corresponding exons, as point mutations are not likely to fully inactivate their regulatory potential. Rather, they need to be removed from the genome by a deletion strategy, similarly to what was previously described in this chapter for noncoding DNA elements such as enhancers and insulators. For example, Han and colleagues generated a large deletion in mice, encompassing the 23 kb of the imprinted lncRNA *Rian* [55]. Interestingly, the authors detected an increase in the expression of genes adjacent to the lncRNA *Rian*, providing evidence of the fact that lncRNA can regulate transcription of nearby genes.

In a study aimed to characterize the function of the lncRNA *Haunt*, Yin and colleagues deployed a number of distinct CRISPR/Cas9-based strategies. To dissect the effect of the *Haunt* transcript from that of the *Haunt* locus, the authors produced a series of deletions of different sizes and directed to functional elements within the element. By deleting small sequences constituting the *Haunt* promoter, resulting in the loss of *Haunt* lncRNA expression, the authors detected an increase in the expression of *Haunt* target genes, suggesting that this transcript has a suppressive role. In contrast, by producing larger deletions, the authors observed a decrease in the expression of the same set of target genes, indicating that the *Haunt* locus acts as a HOXA enhancer. To confirm these data, the authors induced *Haunt* overexpression by knocking in the sequence of a strong constitutive promoter, CAG, just upstream the *Haunt* TSS, using CRISPR/Cas9 to trigger homologous recombination [56]. With this strategy, overexpression of lncRNA could be obtained at the endogenous locus. As miRNAs are often organized into large gene clusters, we could consider that removal of the full cluster could be interesting. In this regard, by injecting two adjacent sgRNA in zebrafish embryos, Xiao and colleagues achieved a very large deletion encompassing a miRNA cluster on zebrafish chromosome 9 [57].

Conclusion

The efficiency, flexibility, and reproducibility of CRISPR/Cas9 approaches triggered many researchers to engage in challenging experiments that were technically very difficult to achieve, or nearly impossible to undertake, just a few years ago. CRISPR/Cas9-mediated genome mutagenesis has provided an outstandingly simple solution to functionally assess both coding and noncoding DNA sequences at their endogenous locations. Currently, with all these target nuclease experimental approaches, the entire mammalian genome can be investigated, in vivo, to decipher the role of coding and noncoding DNA elements in physiology and pathology.

Acknowledgments This review was supported by grants from MINECO projects BIO2012-39980, BIO2015-70978-R, CIBERER (ACCI-2015) and BMBS COST action BM1308 SALAAM to L.M. D.S. was supported by the La Caixa International PhD and EMBO Short Term Fellowship (AST140-2013) programs.

References

1. Cheng Y, Ma Z, Kim BH, Wu W, Cayting P, Boyle AP, Sundaram V, Xing X, Dogan N, Li J, Euskirchen G, Lin S, Lin Y, Visel A, Kawli T, Yang X, Patacsil D, Keller CA, Giardine B; Mouse ENCODE Consortium, Kundaje A, Wang T, Pennacchio LA, Weng Z, Hardison RC, Snyder MP. Principles of regulatory information conservation between mouse and human. Nature. 2014;515(7527):371–5.
2. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009;457(7231):854–8.
3. Schnetz MP, Handoko L, Akhtar-Zaidi B, Bartels CF, Pereira CF, Fisher AG, Adams DJ, Flicek P, Crawford GE, Laframboise T, Tesar P, Wei CL, Scacheri PC. CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. PLoS Genet. 2010;6(7), e1001023.
4. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu YC, Pfenning AR, Wang X, Claussnitzer M, Liu Y, Coarfa C, Harris RA, Shoresh N, Epstein CB, Gjoneska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh KH, Feizi S, Karlic R, Kim AR, Kulkarni A, Li D, Lowdon R, Elliott G, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthall KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Haussler D, Jones SJ, Li W, Marra MA, McManus MT, Sunyaev S, Thomson JA, Tlsty TD, Tsai LH, Wang W, Waterland RA, Zhang MQ, Chadwick LH, Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, Stamatoyannopoulos JA, Wang T, Kellis M. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539): 317–30.
5. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, Diao Y, Liang J, Zhao H, Lobanenkov VV, Ecker JR, Thomson JA, Ren B. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015;518(7539):331–6.
6. Ouyang Z, Zhou Q, Wong WH. ChIP-Seq of transcription factors predicts absolute and differential gene expression in embryonic stem cells. Proc Natl Acad Sci U S A. 2009; 106(51):21521–6.
7. Bauer DE, Kamran SC, Lessard S, Xu J, Fujiwara Y, Lin C, Shao Z, Canver MC, Smith EC, Pinello L, Sabo PJ, Vierstra J, Voit RA, Yuan GC, Porteus MH, Stamatoyannopoulos JA, Lettre G, Orkin SH. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science. 2013;342(6155):253–7.
8. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, Ryan RJ, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015;518(7539):337–43.
9. Lidral AC, Liu H, Bullard SA, Bonde G, Machida J, Visel A, Uribe LM, Li X, Amendt B, Cornell RA. A single nucleotide polymorphism associated with isolated cleft lip and palate, thyroid cancer and hypothyroidism alters the activity of an oral epithelium and thyroid enhancer near FOXE1. Hum Mol Genet. 2015;24(14):3895–907.
10. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gómez-Marín C, Aneas I, Credidio FL, Sobreira DR, Wasserman NF, Lee JH, Puviindran V, Tam D, Shen M, Son JE, Vakili NA, Sung HK, Naranjo S, Acemel RD, Manzanares M, Nagy A, Cox NJ, Hui CC, Gomez-Skarmeta JL, Nóbrega MA. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature. 2014;507(7492):371–5.

11. Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. Genome Res. 2012;22(3):446–55.
12. Guenther CA, Tasic B, Luo L, Bedell MA, Kingsley DM. A molecular basis for classic blond hair color in Europeans. Nat Genet. 2014;46(7):748–52.
13. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet. 2005;6(6):507–12.
14. Bender MA, Bulger M, Close J, Groudine M. Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. Mol Cell. 2000;5(2):387–93.
15. Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, Blow MJ, Cohen JC, Rubin EM, Pennacchio LA. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. Nature. 2010;464(7287):409–12.
16. Volpi SA, Verma-Gaur J, Hassan R, Ju Z, Roa S, Chatterjee S, Werling U, Hou Jr H, Will B, Steidl U, Scharff M, Edelman W, Feeney AJ, Birshtein BK. Germline deletion of Igh 3′ regulatory region elements hs 5, 6, 7 (hs5-7) affects B cell-specific regulation, rearrangement, and insulation of the Igh locus. J Immunol. 2012;188(6):2556–66.
17. Vanden Berghe T, Hulpiau P, Martens L, Vandenbroucke RE, Van Wonterghem E, Perry SW, Bruggeman I, Divert T, Choi SM, Vuylsteke M, Shestopalov VI, Libert C, Vandenabeele P. Passenger mutations confound interpretation of all genetically modified congenic mice. Immunity. 2015;43(1):200–9.
18. Giraldo P, Montoliu L. Size matters: use of YACs, BACs and PACs in transgenic animals. Transgenic Res. 2001;10(2):83–103.
19. Giraldo P, Giménez E, Montoliu L. The use of yeast artificial chromosomes in transgenic animals: expression studies of the tyrosinase gene in transgenic mice. Genet Anal. 1999;15(3-5):175–8.
20. Murisier F, Guichard S, Beermann F. Distinct distal regulatory elements control tyrosinase expression in melanocytes and the retinal pigment epithelium. Dev Biol. 2007;303(2):838–47.
21. Montoliu L, Umland T, Schütz G. A locus control region at −12 kb of the tyrosinase gene. EMBO J. 1996;15(22):6026–34.
22. Giménez E, Giraldo P, Jeffery G, Montoliu L. Variegated expression and delayed retinal pigmentation during development in transgenic mice with a deletion in the locus control region of the tyrosinase gene. Genesis. 2001;30(1):21–5.
23. Seruggia D, Fernández A, Cantero M, Pelczar P, Montoliu L. Functional validation of mouse tyrosinase non-coding regulatory DNA elements by CRISPR-Cas9-mediated mutagenesis. Nucleic Acids Res. 2015;43(10):4855–67.
24. Harms DW, Quadros RM, Seruggia D, Ohtsuka M, Takahashi G, Montoliu L, Gurumurthy CB. Mouse genome editing using the CRISPR/Cas system. Curr Protoc Hum Genet. 2014;83:15.7.1–27.
25. Seruggia D, Montoliu L. The new CRISPR-Cas system: RNA-guided genome engineering to efficiently produce any desired genetic alteration in animals. Transgenic Res. 2014;23(5):707–16.
26. Matys V, Fricke E, Geffers R, Gössling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Münch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, Wingender E. TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. 2003;31:374–8.
27. Li Y, Rivera CM, Ishii H, Jin F, Selvaraj S, Lee AY, Dixon JR, Ren B. CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells. PLoS One. 2014;9(12), e114485.
28. Giraldo P, Martínez A, Regales L, Lavado A, García-Díaz A, Alonso A, Busturia A, Montoliu L. Functional dissection of the mouse tyrosinase locus control region identifies a new putative boundary activity. Nucleic Acids Res. 2003;31(21):6290–305.
29. Giraldo P, Montoliu L. Artificial chromosome transgenesis in pigmentary research. Pigment Cell Res. 2002;15(4):258–64.

30. Canver MC, Bauer DE, Dass A, Yien YY, Chung J, Masuda T, Maeda T, Paw BH, Orkin SH. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. J Biol Chem. 2014;289(31):21312–24.
31. Essletzbichler P, Konopka T, Santoro F, Chen D, Gapp BV, Kralovics R, Brummelkamp TR, Nijman SM, Bürckstümmer T. Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res. 2014;24(12):2059–65.
32. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485(7398):376–80.
33. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326(5950):289–93.
34. Moltó E, Fernández A, Montoliu L. Boundaries in vertebrate genomes: different solutions to adequately insulate gene expression domains. Brief Funct Genomic Proteomic. 2009;8(4): 283–96.
35. Wang J, Vicente-García C, Seruggia D, Moltó E, Fernandez-Miñán A, Neto A, Lee E, Gómez-Skarmeta JL, Montoliu L, Lunyak VV, Jordan IK. MIR retrotransposon sequences provide insulators to the human genome. Proc Natl Acad Sci U S A. 2015;112(32):E4428–37.
36. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, Santos-Simarro F, Gilbert-Dussardier B, Wittler L, Borschiwer M, Haas SA, Osterwalder M, Franke M, Timmermann B, Hecht J, Spielmann M, Visel A, Mundlos S. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell. 2015;161(5):1012–25.
37. Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, van Bueren KL, Chines PS, Narisu N; NISC Comparative Sequencing Program, Black BL, Visel A, Pennacchio LA, Collins FS; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors; NISC Comparative Sequencing Program Authors. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A. 2013;110(44):17921–17926.
38. Hnisz D, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, Bradner JE, Young RA. Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. Mol Cell. 2015;58(2):362–70.
39. Hewitt KJ, Kim DH, Devadas P, Prathibha R, Zuo C, Sanalkumar R, Johnson KD, Kang YA, Kim JS, Dewey CN, Keles S, Bresnick EH. Hematopoietic signaling mechanism revealed from a stem/progenitor cell cistrome. Mol Cell. 2015;59(1):62–74.
40. Canver MC, Smith EC, Sher F, Pinello L, Sanjana NE, Shalem O, Chen DD, Schupp PG, Vinjamur DS, Garcia SP, Luc S, Kurita R, Nakamura Y, Fujiwara Y, Maeda T, Yuan GC, Zhang F, Orkin SH, Bauer DE. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature. 2015;527(7577):192–7.
41. Chiba K, Johnson JZ, Vogan JM, Wagner T, Boyle JM, Hockemeyer D. Cancer-associated TERT promoter mutations abrogate telomerase silencing. Elife. 2015;4. doi:10.7554/eLife.07918.
42. Katainen R, Dave K, Pitkänen E, Palin K, Kivioja T, Välimäki N, Gylfe AE, Ristolainen H, Hänninen UA, Cajuso T, Kondelin J, Tanskanen T, Mecklin JP, Järvinen H, Renkonen-Sinisalo L, Lepistö A, Kaasinen E, Kilpivaara O, Tuupanen S, Enge M, Taipale J, Aaltonen LA. CTCF/cohesin-binding sites are frequently mutated in cancer. Nat Genet. 2015;47(7):818–21.
43. Mansour MR, Abraham BJ, Anders L, Berezovskaya A, Gutierrez A, Durbin AD, Etchin J, Lawton L, Sallan SE, Silverman LB, Loh ML, Hunger SP, Sanda T, Young RA, Look AT. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science. 2014;346(6215):1373–7.

44. Puente XS, Beà S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, Munar M, Rubio-Pérez C, Jares P, Aymerich M, Baumann T, Beekman R, Belver L, Carrio A, Castellano G, Clot G, Colado E, Colomer D, Costa D, Delgado J, Enjuanes A, Estivill X, Ferrando AA, Gelpí JL, González B, González S, González M, Gut M, Hernández-Rivas JM, López-Guerra M, Martín-García D, Navarro A, Nicolás P, Orozco M, Payer ÁR, Pinyol M, Pisano DG, Puente DA, Queirós AC, Quesada V, Romeo-Casabona CM, Royo C, Royo R, Rozman M, Russiñol N, Salaverría I, Stamatopoulos K, Stunnenberg HG, Tamborero D, Terol MJ, Valencia A, López-Bigas N, Torrents D, Gut I, López-Guillermo A, López-Otín C, Campo E. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015;526(7574):519–524.
45. Zhang T, Cooper S, Brockdorff N. The interplay of histone modifications—writers that read. EMBO Rep. 2015;16(11):1467–81.
46. Ahuja N, Sharma AR, Baylin SB. Epigenetic therapeutics: a new weapon in the war against cancer. Annu Rev Med. 2016;67:73–89.
47. Falahi F, van Kruchten M, Martinet N, Hospers GA, Rots MG. Current and upcoming approaches to exploit the reversibility of epigenetic mutations in breast cancer. Breast Cancer Res. 2014;16(4):412.
48. Maes T, Carceller E, Salas J, Ortega A, Buesa C. Advances in the development of histone lysine demethylase inhibitors. Curr Opin Pharmacol. 2015;23:52–60.
49. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, Costello JF, Wilkinson MF, Joung JK. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat Biotechnol. 2013;31(12):1137–42.
50. Kearns NA, Pham H, Tabak B, Genga RM, Silverstein NJ, Garber M, Maehr R. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nat Methods. 2015;12(5):401–3.
51. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol. 2015;33(5):510–7.
52. Kungulovski G, Jeltsch A. Epigenome editing: state of the art, concepts, and perspectives. Trends Genet. 2015; pii: S0168-9525(15)00212-7.
53. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell. 2013; 154(1):26–46.
54. Xie C, Yuan J, Li H, Li M, Zhao G, Bu D, Zhu W, Wu W, Chen R, Zhao Y. NONCODEv4: exploring the world of long non-coding RNA genes. Nucleic Acids Res. 2014;42(Database issue):D98–103.
55. Han J, Zhang J, Chen L, Shen B, Zhou J, Hu B, Du Y, Tate PH, Huang X, Zhang W. Efficient in vivo deletion of a large imprinted lncRNA by CRISPR/Cas9. RNA Biol. 2014;11(7):829–35.
56. Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, Zhao Y, Shen B, Huang X, Zhu H, Orkin SH, Shen X. Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. Cell Stem Cell. 2015;16(5):504–16.
57. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, Zu Y, Li W, Huang P, Tong X, Zhu Z, Lin S, Zhang B. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res. 2013;41(14), e141.

At the Conflux of Human Genome Engineering and Induced Pluripotency

Knut Woltjen, Fabian Oceguera-Yanez, Harunobu Kagawa, and Shin-Il Kim

Abstract Human induced pluripotent stem cells (iPSCs) represent a personalized stem cell source and enable research using the human as a model genetic system. Although iPSCs have been available for nearly a decade, simple application of efficient genome modification—a mainstay of genetics in the long-used mouse model—has only recently come to fruition. Recombinant and programmable nucleases induce targeted DNA damage and exploit native DNA-repair machinery to generate random mutations or designer modifications through a template-mediated process. In this review, we provide an overview of state-of-the-art nuclease technologies such as ZFN, TALEN, and CRISPR/Cas9 and their utility for genome engineering of human iPSCs. We explore how nucleases may be used to edit the genome with base-pair precision, and methods for the detection and avoidance of off-target cleavage. Finally, we highlight sources of genetic and technical variation in iPSCs, and propose resolutions to the question of appropriate isogenic controls.

Keywords Human induced pluripotent stem cell • iPSC • Gene targeting • ZFN • TALEN • CRISPR/Cas9 • Nuclease • Disease model • Gene correction • Isogenic controls

Abbreviations

AAVS1	Adeno-associated virus integration site 1
CRISPR	Clustered regularly interspaced short palindromic repeat
crRNA	CRISPR RNA
ddPCR	Droplet digital PCR

K. Woltjen (✉)
Center for iPS Cell Research and Application (CiRA), Kyoto University,
53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan

Hakubi Center for Advanced Research, Kyoto University, Kyoto 606-8501, Japan
e-mail: woltjen@cira.kyoto-u.ac.jp

F. Oceguera-Yanez • H. Kagawa • S.-I. Kim
Center for iPS Cell Research and Application (CiRA), Kyoto University,
53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_3

DSB	Double-strand break
DSBR	Double-strand break repair
ESC	Embryonic stem cell
GFP	Green fluorescent protein
HDR	Homology-directed repair
HLA	Histocompatibility leukocyte antigen
indel	Insertion or deletion
iPSC	Induced pluripotent stem cell
ITR	Inverted terminal repeat
IVF	In vitro fertilization
MMEJ	Microhomology-mediated end-joining
NHEJ	Nonhomologous end-joining
PAM	Protospacer adjacent motif
PB	*piggyBac* transposon
PBase	*piggyBac* transposase
PCR	Polymerase chain reaction
PSC	Pluripotent stem cell
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
ROCKi	Rho-kinase inhibitor
sgRNA	Single guide RNA
SNV	Single nucleotide variation
ssODN	Single-strand oligonucleotide
T7E1	Bacteriophage T7 endonuclease I
TALEN	Transcription activator-like effector nuclease
tracrRNA	Trans-acting crRNA
WGS	Whole-genome sequencing
ZFN	Zinc-finger nuclease

Induced Pluripotency and the Human Genetic Model Organism In Vitro

From the inner cell mass (ICM) of fertilized embryos, James Thomson first derived human embryonic stem cells (ESCs) [1]. These novel cells have two key properties: first, they are capable of indefinite cell division in culture (self-renewal), and second, as do their biological counterparts, they maintain the capacity to differentiate into all cells and tissues of the embryo and adult (pluripotency). The unchallenged advantage of human ESCs over other experimental cell systems has been this capacity for differentiation, either in vivo via teratoma [2] or in vitro via adherent or three-dimensional (3D) cell culture [3, 4]. An application in disease research and regenerative medicine for ESCs was immediately apparent; however, the embryonic source of material has remained an ethical controversy [5].

One decade has now passed since Kazutoshi Takahashi and Shinya Yamanaka first demonstrated that mouse somatic cells could be reverted through the expression of four transcription factors, Oct3/4, Sox2, Klf4, and c-Myc, to a primitive embryonic-like stem cell state [6]. Derivation of induced pluripotent stem cells (iPSCs) from human somatic cells (Fig. 1) followed shortly thereafter [7], marking a partial ethical resolution and profound technical contribution [8]. Compared to ESCs, iPSCs present an additional benefit for disease modeling and putative therapies: derived from a consenting individual, they represent personalized stem cells. Moreover, in contrast to ESCs from terminated embryos, iPSCs may be linked to the health and well-being of a living person, complemented by a recorded lifetime medical history. Thus, combined with in vitro differentiation to cells and tissues (Fig. 1), human iPSCs present a proxy by which individualized genetic variation may be accessed to understand the relevance to personal health [9].

Pluripotent stem cells (PSCs)—whether ESCs derived from the human embryo or iPSCs derived through reprogramming—display key properties of a tractable genetic system: a short generation time (~15 h), a high proportion of cells in S-phase [10], indefinite proliferation, ease of culture, a propensity for DNA transduction, and selection by antibiotics or genetic complementation followed by clonal isolation and expansion. With more recent advances such as Rho-kinase inhibition for improved single-cell survival [11], and feeder-free cell culture methods using defined matrices and media [12, 13], human PSC handling is more akin to murine PSC counterparts by means of single cell passage, and high-throughput 96-well clonal maintenance and expansion [14].

As is explored in the following sections, the marriage of iPSC technology with a new generation of genetic engineering tools has enabled the precise transfer of

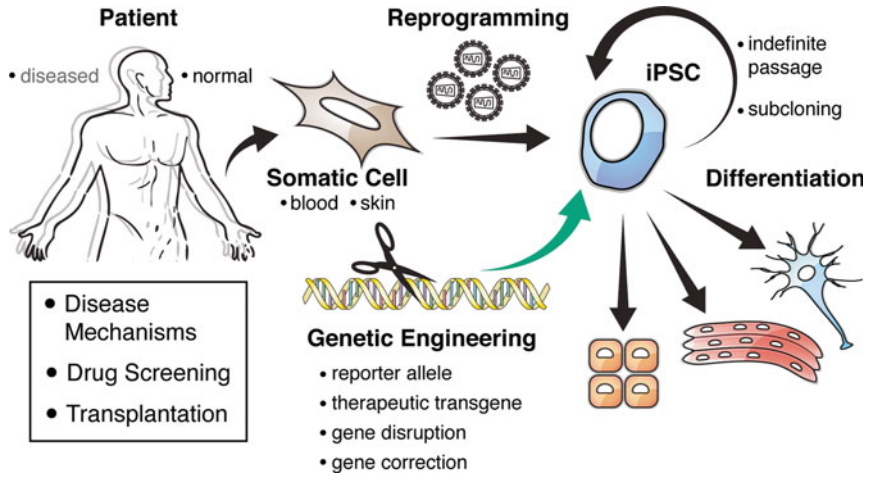

Fig. 1 The source, characteristics, and applications of human pluripotent stem cells. Genome engineering is applied to generate or validate human models of development and disease

reporter or therapeutic transgenes, gene disruption, or even gene correction (Fig. 1). It is remarkable to reflect on the speed and relative ease at which both iPSC and genome engineering technologies have been adapted and merged for in vitro disease modeling and drug screening.

Rise of the Genome Editing Machines

Genetic manipulation by gene targeting is a mainstay of functional genomics. Those fundamental principles of gene targeting first outlined by Mario Cappechi using positive-negative selection in mouse ESCs [15, 16] are duly applicable to human PSCs. However, even following these guidelines, the first gene-targeting experiments in human ESCs [17] indicated that gene-targeting rates would be typically lower than observed in the mouse, leaving an obvious need for improvement.

The formation of double-strand breaks (DSBs) in genomic DNA occurs naturally during DNA replication or in response to stresses such as ionizing radiation, and are vital to resolve recombination during meiosis and the production of immune system diversity [18]. DSB repair (DSBR) by end-resection and nonhomologous end-joining (NHEJ) can be inherently mutagenic, whereas homology-directed repair (HDR) can faithfully restore DNA sequence in the presence of a template donor DNA (such as the sister chromatid). It was therefore hypothesized that intentional formation of DSBs at target loci could enhance gene-targeting frequencies via a custom donor DNA [19, 20]. The demonstration that the FokI nuclease domain is separable from DNA-binding domains [21] suggested a method by which nucleases could be engineered with novel specificity.

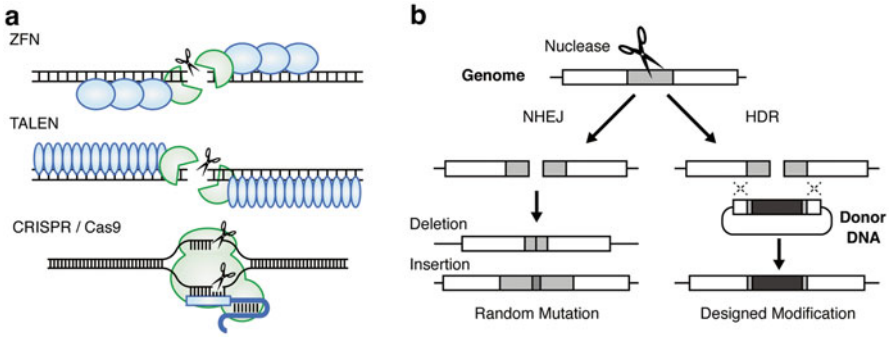

Fig. 2 Common nuclease systems used to stimulate double-strand break repair (DSBR) at target genomic sites. (**a**) Zinc-finger nucleases (ZFN) and TAL effector nucleases (TALEN) are composed of dimeric nuclease domains addressed by engineered DNA-binding domains. The Cas9 nuclease is addressed by a synthetic guide RNA molecule. Endonuclease components are *green*; targeting components are *blue*. (**b**) DSBs recruit endogenous repair machinery allowing genetic modification by nonhomologous end-joining (NHEJ) or template-mediated homology-directed repair (HDR) pathways

Zinc-finger nucleases (ZFNs; Fig. 2a, top) built a foundation for recombinant endonuclease applications to enhance gene targeting through targeted DSBR, and remain a powerful tool in genome engineering even today [22]. Composed of a DNA-binding domain-encoding specificity and a FokI nuclease domain, ZFNs function as paired proteins that position and dimerize FokI monomers to cleave at a target locus [23]. The binding of zinc fingers to DNA triplets is modular [24]; however, finger–triplet interaction properties have been shown to be highly context and neighbor dependent, such that engineering custom ZFNs remains notoriously difficult. This problem has been partially addressed using validated libraries of submodules composed of two or three fingers [25]; however, screening for functional ZFNs is a resource-heavy endeavor.

TAL effector nucleases (TALENs; Fig. 2a, middle) broke the barrier between technical novelty and practical application [26], priming research laboratories through in-house nuclease design and production [27]. Plant pathogenic *Xanthomonas* spp. secrete TAL Effector (TALE) proteins, which activate host gene expression, resulting in a metabolic advantage to the invader [28]. The TALEs represent a unique class of proteins that bind DNA in a 1:1 modality [29, 30], making engineered design of TALENs more straightforward than that of ZFNs. The nature of this protein–DNA interaction is mediated through polymorphic protein repeats that display little degeneracy and no obvious neighbor effects. In a large-scale in vivo screen in zebrafish, TALENs were found to be more mutagenic than ZFNs [31]. Presumably, the increased tolerance of the spacer region provides a larger substrate for exonuclease activity, resulting in broad deletions compared to the conservative resection observed using tightly juxtaposed ZFNs.

CRISPR/Cas9 (Fig. 2a, bottom) has stolen the proverbial 'show,' capturing the attention of academia and public alike as it rapidly transcended from discovery as the hunter–killer of a potent anti-phage adaptive bacterial immune system [32], to experimental modulation and design [33] for genome engineering purposes [34]. The Cas9 protein, a general endonuclease that produces DSBs through HNH and RuvC nuclease domains, forms a ribonucleoprotein (RNP) complex with bacterially processed short CRISPR RNAs (crRNA) and trans-acting crRNA (tracrRNA) that pair with foreign genomic DNA targets to address and activate nuclease activity [35]. Biochemical characterization of the key CRISPR components, by Jennifer Doudna and Emmanuelle Charpentier [33], indicated that programmable cleavage could be achieved through the custom design of a hybrid crRNA–tracrRNA single guide RNA molecule (sgRNA), which could be simply co-expressed or transfected as RNA along with the *Streptococcus pyogenes* Cas9 (SpCas9) protein. Thus, the CRISPR/Cas9 system could theoretically be programmed to cleave any 20-nt sequence upstream of a 5′-NGG-3′ protospacer adjacent motif (PAM). This report was immediately followed by back-to-back proof-of principle experiments describing genome engineering in human cells [36, 37]. During the past 3 years, CRISPR/Cas9 technology has enabled gene knockouts across previously inaccessible genetic model organisms [38] and high-throughput genomic screens [39, 40], highlighting the simplicity of design and ease of application.

SpCas9 is by far the most commonly used CRISPR system. However, active variants from other bacterial species such as *Staphylococcus aureus* (SaCas9, 5′-NNGRRN-3′ PAM) [41] and *Neisseria meningitidis* (NmCas9, 5′-NNNNGATT-3′ PAM) [42] have been applied for genome editing with variable success. As proteins of bacterial origin, rational design of Cas9 in prokaryotes is conventional, leading to new variants of SaCas9 with modified PAM specificity, and therefore broader targeting ranges [43]. Mining prokaryotic genome databases through homology, cloning, and functional validation has yielded family members with new properties. Differing from SpCas9, *Francisella novicida* Cpf1 (FnCpf1) requires only a single guide RNA, and recognizes a 5′-TTN-3′-PAM, therefore accessing completely different sequence space [44]. The rich diversity of CRISPR systems in bacteria suggests that additional nucleases with distinct properties remain to be discovered.

Biochemical subtleties of DNA recognition and cleavage aside, engineered nucleases enhance random mutagenesis and gene targeting by eliciting endogenous DSBR pathways (Fig. 2b). As the NHEJ mutation spectrum is essentially random, it provides allelic depth for clonal cell panels, yet can complicate high throughput screening [45]. Under special circumstances of genomic sequence context and DSB position, DSBR can be driven by subtle regions of microhomology to produce indels in a predictable manner [46]. Bi-allelic DSBs can allow for homozygous targeting by HDR, an event achieved rarely with classic gene targeting [47]. Combinatorial events, such as HDR-mediated targeting of one allele, and NHEJ knockout of the other, can be a boon or a bane. Because of the promiscuity of the Cas9 protein, combinatorial approaches using multiplexed sgRNAs have led to multiple mutations in a mouse stem cell or embryo [48, 49], accelerating the analysis of multiple genetic interactions and emphasizing the power of the CRISPR/Cas9 nuclease system for functional genomics studies.

Adding Function to iPSCs Through Transgenesis at Safe-Harbor Loci

Gene targeting may be used to eliminate or alter endogenous genes, or introduce new functions. Transgenesis with viral or transposon systems has the advantage of being robust and rapid [14], yet as a trade-off does not directly control for integration site and therefore requires the use of populations or screening multiple clones to discern suitable or comparable gene expression levels. Nucleases permit transgenes to be introduced into defined loci and therefore minimize clonal variation by moderating position effects [50]. HDR-targeted transgenesis includes applications such as cDNA rescue of mutant genes and fluorescent knock-in alleles to report endogenous gene expression or simply label cells constitutively [51].

Perhaps the most well known “safe-harbor” locus is AAVS1, a hotspot for adeno-associated virus insertion located within intron 1 of the PPP1R12C gene [52]. AAVS1 is akin to the mouse ROSA26 locus [53], providing a reliable transgene

expression with no known phenotype resulting from homozygous transgene insertions [47]. Targeting and expression of cDNAs from the AAVS1 locus has rescued monogenic diseases such as X-linked chronic granulomatous [54] and α-thalassemia [55]. Conversely, overexpression of dominant negative ion channel genes KCNQ1 and KCNH2 from the AAVS1 locus can recapitulate Long-QT syndrome for the development of an isogenic in vitro drug-screening platform [56].

Other safe harbors, such as the X-linked hypoxanthine phosphoribosyltransferase 1 (HPRT1) locus, is permissive for constitutive expression [57], yet disruption causes HPRT1-deficiency spectrum diseases ranging from gout to Lesch–Nyhan syndrome. The human L-gulono-γ-lactone oxidase (GULOP) locus is a nonfunctional pseudogene in humans [58, 59] presumed to avoid phenotypic effects. Yet, transgene expression in pluripotent and differentiated lineages is less well described. Beyond gene disruption, the local effects by potent transgenic promoters on endogenous gene expression must also be considered [60]. One such example is the citrate lyase beta-like (CLYBL) locus that lies in a gene-deficient region of human chromosome 13 and claims to confer less severe effects on local gene expression [61]. Finally, in a mouse model of hemophilia A and B, expression of human factors VIII and IX from the endogenous albumin locus achieved long-term expression of transgenes at therapeutic levels [62]. Therefore, context-dependent safe harbors may be found in loci that are active in the target-differentiated cell type yet repressed in others, and not associated with a known haploinsufficiency phenotype.

Achieving Seamless Genome Engineering for Accurate Disease Models

In the interest of generating faithful models of human genetic disease, engineered changes that recapitulate single-nucleotide variations (SNVs) would be preferred over crude knockouts. The de facto test for evaluating the role of candidate mutations in disease is to repair the mutation in patient iPSCs, or to recapitulate it in otherwise normal iPSCs [63]; true correction or recreation of patient-specific mutations would require approaches that are free of residual foreign genetic elements.

Classic gene targeting [16] deposits antibiotic-positive selection cassettes to enrich for HDR-mediated events (Fig. 3a). Retention of such elements is invaluable for producing knockouts, and reconcilable with the integration of reporters [47] or even therapeutic transgenes [64]. However, in the interest of modifying small regions of DNA—or in the extreme case, single nucleotides—selection cassettes and other elements can disrupt the native locus and may even cause unpredictable pleiotropic effects [65]. Removal of antibiotic selection cassettes is typically performed through site-specific recombinase-mediated excision [66]. In this approach, the recognition sites for Cre (*loxP*), Flp (FRT) recombinases [67] flank the selection cassette, which is introduced juxtaposed to the mutation (Fig. 3b, left). Following the selection of targeted clones, the cassette is excised by transient recombinase expression, yet a nontrivial single recombinase site (34 bp in the case of *loxP*) remains. Although residual elements may

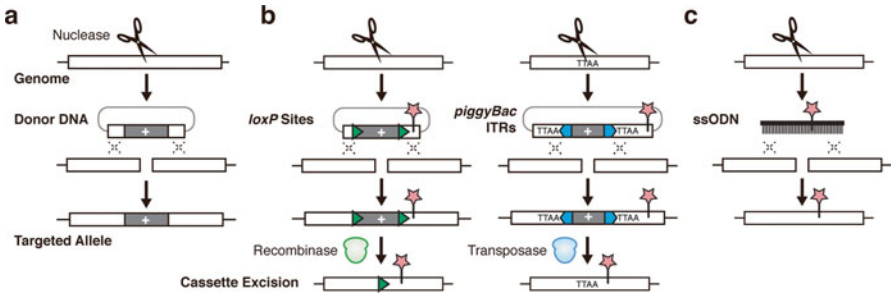

Fig. 3 Derivation of human induced pluripotent stem cells (iPSCs) engineered by HDR. (**a**) Classic gene targeting events are enriched by positive selection. (**b**) Excision of antibiotic selection markers by Cre recombinase (left) leaves behind *loxP* sites, while PB transposase (right) removes cassettes seamlessly from endogenous or engineered TTAA tetranucleotides to deposit point mutations. (**c**) Mutation deposition by short single-strand oligonucleotides (ssODNs) obviates the need for excision, yet requires intensive screening

be positioned in "neutral" genetic regions such as introns, unexpected effects on gene expression and the predicted phenotype remain probable [64].

As an alternative to recombinases, the *piggyBac* (PB) transposon undergoes high-fidelity seamless excision from mouse and human iPSCs [68], and has been developed as an excisable positive/negative selection cassette for genome modification [69]. One caveat is that PB elements excise only from TTAA tetranucleotides, such that a TTAA must be present, or silently engineered near the mutation (Fig. 3b, right). PB provides more flexibility and a subtler footprint than recombinases, yet excision frequencies are locus dependent, and reintegration of the transposon may occur stochastically, whereas excision-prone transposase variants [70] may display higher rates of mutagenesis.

Diverging from classic targeting vector-based genome modification relying on antibiotic enrichment, short single-strand oligonucleotide (ssODN) templates have been employed in combination with ZFNs [71, 72], TALENs [73], and CRISPR/Cas9 [48]. In this approach, ssODNs typically more than 100 nt in length carry sufficient homology to deposit point mutations into nuclease-cleaved loci in a single step without codeposition of foreign sequences (Fig. 3c), providing a clear advantage over recombinase-based methods [74]. It should be noted that ssODN-modified loci that retain the nuclease target site are potentially subject to recleavage and mutagenic NHEJ repair. Silent mutations that prevent nuclease recognition and recleavage detract from the subtlety of the method, but may be necessary to avoid additional screening. Moreover, aberrant ssODN insertions at on- or off-target sites [75] or random mutations on-target [76] may occur under normal conditions or as a reflection of oligo quality and are extremely difficult to predict and detect. Although ssODN-mediated targeting events are frequent in cell lines, the low frequency of correct targeting in iPSCs (>1 %) [48, 72, 73], compounded with possible mutagenic events, demands robust and sophisticated selection.

One advanced approach to detect correct gene editing employs serial population screening, a type of sib-selection for human iPSCs where mutation-containing populations are monitored by droplet digital polymerase chain reaction (ddPCR) and enriched using serial sub-fractionation [77]. In developing this technique, the authors successfully deposited mutations into five disease-associated genes (*PHOX2B*, *PKP2*, *RBM20*, *PRKAG2*, and *BAG3*). Population screening by ddPCR is robust but not trivial, requiring custom TaqMan assays, sophisticated instrumentation, and additional iPSC passages. A streamlined approach to derive gene-corrected iPSCs that combined CRISPR/Cas9 gene targeting with the somatic cell reprogramming process [76], reported gene knock-in efficiencies as high as 5 %, and ssODN-mediated gene correction rates as high as 8 %. Although useful during de novo iPSC derivation, this approach is obviously not applicable to previously established iPSC lines. Finally, frequencies of desirable targeting using ssODNs may still see improvements through lessons learned from the biochemistry of DNA opening and Cas9 cleavage. As the sgRNA nontarget (unbound) strand is released first, ssODNs with positioning and complementarity to the nontarget DNA strand are more effective at inducing HDR, up to 60 % in HEK293T cells [78]. Applications of these findings in iPSCs hold promise.

Avoiding Unwanted Outcomes: Off-Target Cleavage and Mosaicism

Nuclease cleavage of the genome is by no means infallible, and undesirable DSBR events may occur through surreptitious cleavage at sites other than the chosen target region. In these cases, DSBs repaired preferentially through NHEJ may result in subtle indels (Fig. 2b) with no capacity for counterselection. Such "off-target" effects may or may not have phenotypic consequences.

Unbiased off-target detection using whole-genome sequencing (WGS) can evaluate genome-engineered iPSC clones [79, 80], yet the depth of data and the threshold for detecting rare mutations argue against the practicality of the approach. Exome sequencing simplifies analysis, yet provides data for only a small portion of the genome. Targeted screening methods based on degenerate sequence similarity between the sgRNA and nontarget regions of the host genome provide an off-target candidate list that may be verified using conventional NHEJ detection methods such as the T7E1 hybrid-cleavage assay [81], Sanger sequencing with decomposition [82], or deep sequencing of amplified products [46]. However, these biased approaches are time consuming and limited by the quality of prediction algorithms for candidate off-target sites.

Off-target screens relying on the functional properties of nucleases have the potential to focus screening efforts without user bias. Chromatin immunoprecipitation using Cas9 antibodies [83, 84] can detect sites of Cas9 interaction with the genome but are not related directly to DNA-cleavage events. Linear amplification-mediated high-throughput, genome-wide, translocation sequencing (LAM-PCR HTGTS) is a cumulative method that detects off-target cleavage by virtue of

genomic translocations formed between nuclease-generated or even endogenous DSBs [85], indicating a two-break-minimum detection limit. On the other hand, single NHEJ events have been shown to capture foreign DNA elements such as integration defective lentiviral vectors (IDLV) [86, 87], which can then act as tags for targeted sequencing efforts. GUIDE-seq applies this same principle, yet uses oligonucleotide tags compatible with next-generation sequencing to streamline sample processing and data integration [88]. BLESS (direct in situ breaks labeling, enrichment on streptavidin, and next-generation sequencing) attempts to capture a snapshot of the fragmented genome within cells, but requires complex fixation and manipulation steps [89]. DiGenome sequencing is an in vitro approach to genomic DSB detection using WGS to detect indels as DNA fragment ends [90]. Differences in detection profiles for these methods may reflect the methodology and must ultimately be verified experimentally.

Refining the detection of off-target cleavage is a crucial endeavor, yet does not directly prevent the causative insult to the genome. Therefore, it would be prudent to develop engineering methods that minimize off-target cleavage events or increase on-target cleavage specificity. One straightforward approach could be to temporally limit the expression of Cas9 and sgRNAs. However, simply reducing the amount of expression vector DNA transfected does not reduce the relative rates of off-target cleavage [91]. In contrast to plasmids, which can express over periods of 3 to 4 days or even integrate randomly into the genome, delivery as in vitro transcribed (IVT) mRNA limits the nuclease expression window to 1 to 2 days and yet is still effective for on-target cleavage. An additional step toward restricted nuclease activity is to produce RNP particles through the in vitro combination of commercially available recombinant SpCas9 protein and IVT or synthetic sgRNAs, followed by delivery directly into iPSCs by electroporation or chemical transfection [92, 93]. An in-depth analysis of the off-target outcomes from such procedural changes is pending.

It is clear, however, that limiting nuclease activity temporally has the potential to reduce mosaicism under conditions normally presumed to produce clonal iPSCs. Mosaicism can arise from unique DNA cleavage and DSBR events in the daughter cells of nuclease-transfected iPSCs, resulting in two or more divergent populations in a drug-selected colony [94]. Mosaicism confounds the detection of off-target effects, which may be present below the threshold of detection in the total iPSC population. Interestingly, sib-selection procedures involving rounds of serial subcloning from the starting population [77] impose a temporal separation of nuclease treatment and physical cloning events to derive truly clonal iPSC populations.

Engineering native nuclease behavior to reduce or prevent off-target cleavage was initially proposed for recombinant FokI nuclease domains [95]. By inactivating the catalytic domain of one monomer in a ZFN dimer, ZFNickases were shown to have lower levels of off-target mutagenesis, albeit with an overall reduction in on-target HDR activity [96]. Similarly, a derivative of SpCas9 in which the RuvC nuclease domain has been inactivated by mutagenesis (SpCas9n, D10A) acts as a DNA nickase [37]. This hobbled enzyme has been used in juxtaposed pairs to produce staggered nicks, and touted as having lower off-target cleavage activity than their full active counterparts because rogue binding of a SpCas9n monomer would produce single-strand

nicks rather than DSB [97]. Yet, it is important to remember that nicked DNA intermediates can still be processed by NHEJ mechanisms, resulting in an off-target indel [98], suggesting that alternative approaches still require consideration.

With solution of the DNA/RNA hybrid-bound SpCas9 protein structure [99, 100] came the possibility of rational SpCas9 engineering. Modeling revealed a positively charged groove between the nuclease and PAM interacting domains, proposed to be involved in stabilizing the nontarget DNA strand. Mutagenesis of K848A, K1003A, and R1060A residues within the groove retained approximately 60 % on-target activity, while increasing sensitivity to sgRNA mismatches, most notably outside the 7–12 nt "seed" region [44]. In another approach, diminished bonding energy through quadruple mutagenesis of DNA-contacting N497A, R661A, Q695A, and Q926A residues produced a high-fidelity variant of SpCas9 (SpCas9-HF1) with undetectable off-target activity [101]. On-target activity was reported to be 70 % of the native SpCas9, a modest compromise for higher specificity.

Modulation of the RNA component of the Cas9 RNP complex has also been shown to positively affect on- and off-target cleavage ratios. It was suggested that truncated sgRNAs (truRNAs) may gain cleavage specificity as a trade-off for activity [102], yet 16-nt versus the standard 20-nt sgRNA molecules has not become a norm for CRISPR experiments. Optimal sgRNA design has been shown to affect on-target cleavage activity [103]. More recently, revision of the rule set governing sgRNA design by Doench and colleagues suggests a predictive scoring system for increasing on-target activity while avoiding off-target cleavage, as demonstrated using a genome-wide knockout screen [104]. It remains to be seen how the community at large will adopt these bioinformatic rule sets, and if they hold true in various experimental situations.

Selection of Isogenic Clones and Technical Controls

Reprogramming technology captures the genome of the patient as a pluripotent cell resource, enabling in vitro modeling of disease that, by necessity, separates the cell from the patient. Differing from animal models, phenotyping results are therefore limited by the sophistication of in vitro cellular differentiation [105] and assay evaluation criteria. To directly link genotypes to phenotypes, appropriate control cell lines are of utmost importance.

Control iPSC lines represent a selected or engineered group of iPSCs that are genetically matched for the purpose of excluding erroneous variation and increasing the accuracy of disease studies [106]. iPSCs from unrelated normal individuals have been used to produce target cell disease controls [107], taking into account that their genetic backgrounds may vary by degrees (Fig. 4a). As such, the number of unrelated iPSC clones that must be analyzed in parallel to define a genotype–phenotype correlation increases in relation to the statistical power required (Fig. 4b). Within practical limits, the required number of control iPSC lines depends mainly on the strength and correlation of the in vitro phenotype with clinical presentation and the

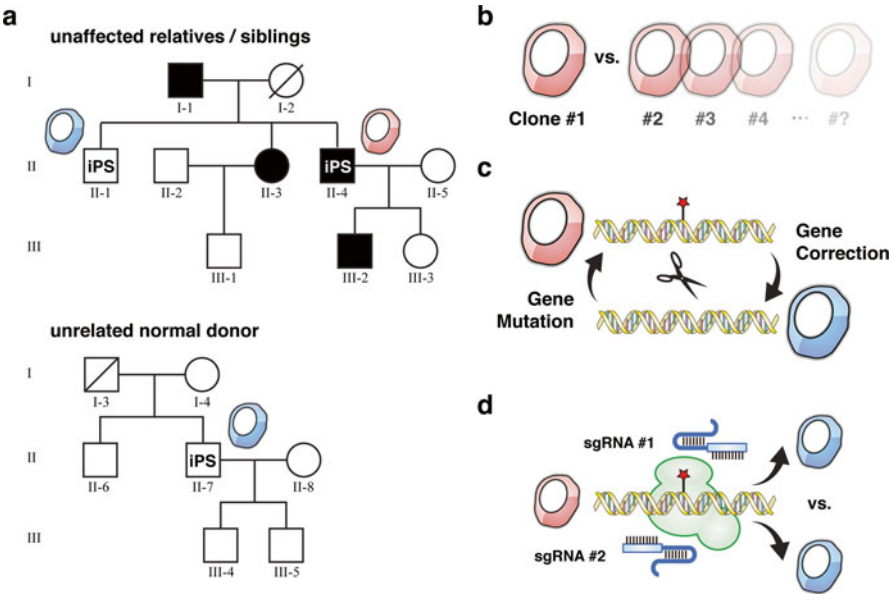

Fig. 4 Appropriate sources of isogenic control iPSC clones. (**a**) iPSCs from unaffected siblings or normal donors are typically used as controls. (**b**) Multiple iPSCs may be used to reduce noise from clonal variation. (**c**) True isogenic controls may be produced through genome engineering. (**d**) To preclude phenotypic effects from off-target cleavage, different sgRNAs may be used to produce the same genomic modification

complex influence of genetic background variation on phenotypes observed within the patient population [108]. One standard for reducing genetic variation has been to compare disease iPSCs to normal iPSCs from unaffected siblings who share much of their genetic background with the affected donor by blood relationship (Fig. 4a). Intriguingly, as ESCs are often derived from pools of discarded in vitro fertilization (IVF) material [1], there may in fact be a higher rate of sibling relationship among publically available ESCs than iPSC lines. However, potential racial bias and an association with a higher incidence of infertility-related alleles, along with a reported marked difference in differentiation capacity between ESC lines [109], may further offset this proposed benefit. On the other hand, the documented medical background of the donor combined with a deep genetic analysis may help predict the severity of phenotypic deviation between experimental and control iPSC lines.

Subtle differences in the genomes and epigenomes between iPSC lines may influence in vitro phenotypes [110]. Concerns over the accumulation of mutations throughout the reprogramming process as a result of proliferative stress have been raised [111]. Conversely, more recent studies have shown that iPSC derivation is inherently stable at the genetic level [112], suggesting that the risk of genetic drift arises during extended in vitro culture and is therefore similar for both ESCs and iPSCs. However, the process of iPSC derivation itself is selective, such that preexisting

somatic mutations in the patient's donor tissue can be clonally amplified [113, 114]. It has also been proposed that reprogrammed cells might retain an epigenetic memory of their somatic source, which could influence differentiation capacity [115]. Interestingly, such epigenetic memory has been disputed by the observation that more significant variation in differentiation capacity occurs as a result of genetic background than somatic tissue source [116, 117]. Still, these uncertainties regarding inherent and acquired phenotypic variation strongly argue the case for isogenicity.

Fortunately, iPSCs themselves are inherently isogenic with their donor, and through the application of subtle nuclease-mediated genome-editing approaches described above, gene-corrected iPSCs can be derived directly from donor iPSCs (Fig. 4c) [63]. Similarly, well-characterized normal iPSCs will retain isogenicity if converted to diseased iPSCs using nuclease techniques. When patient-specific iPSCs cannot be procured, recreating mutations by genome editing provides a novel material for the study of genetic effects on disease progression and severity in a defined genetic background. Quality-controlled normal iPSCs could be accessed from one of many proposed stem cell "libraries," which aim to generate clinical-grade and HLA haplotype-matched control iPSC lines for therapeutic applications [118]. However, it should be cautioned that in the conversion of normal iPSCs into diseased iPSCs, disease phenotypes might be masked by protective alleles. Candidate gene disruption in multiple ethnic backgrounds may therefore be necessary to exclude complex genetic effects [119].

Phenotypic variations between experimental iPSC lines and isogenic controls may have a technical origin. Off-target nuclease effects require labor-intensive screening to detect and might contribute to the observed phenotype. As an alternative to deep sequencing or comparing multiple gene-corrected clones from a single experiment (Fig. 4b, c), it is advised to instead make use of a second sgRNA with its own distinct off-target profile (Fig. 4d). In this way, it is possible to rule out common off-target events between separately derived clones as a direct influence on phenotype. Similarly, employing PB-mediated gene targeting and excision for precise editing [120], reintegration of the transposon may occur stochastically. Yet, these clones may still prove useful for validating phenotypes, because each clone should represent a novel reintegration event. Splinkerette PCR-based methods for mapping reintegrations [121] could help predict the influence on genomic integrity and possible phenotypic changes. Considering these sources of technical variation and their logical solutions, genome engineering (Fig. 4c) stands as the strictest method to maintain isogenicity within control iPSCs.

Conclusions

The combination of iPSC and nuclease technologies, particularly CRISPR/Cas9, has generated a true paradigm shift in modeling human genetics and disease. Although more accessible than ESCs, patient-specific iPSCs still require informed consent, and can prove to be morally and monetarily extravagant research materials.

Applying genome-editing technologies, there is no longer a need to initiate disease modeling with the procurement of patient-specific iPSCs. Instead, candidate mutations or allelic series may be first engineered singly or in combinatorial fashion into genetically and phenotypically defined "reference" ESCs or iPSCs. Once available, the panel of iPSC "standards" may be used to refine in vitro physiological assays. Finally, as required, patient-specific iPSCs may be screened using the optimized assay system to interrogate candidate mutations and the effect of native genetic background. Future avenues of research will most certainly entail combined gene editing and reprogramming strategies, bringing to fruition both preclinical and clinical applications of stem cell technology to personalized medicine.

Acknowledgments We appreciate personal communications with Dana Carroll, Bruce R. Conklin, Gregory D. Davis, J. Keith Joung, Jin-Soo Kim, and William C. Skarnes, as well as the important contributions to the field by the many scientists whose research is not cited here for the sake of brevity. This work was funded in part by grants to K.W. from the Japan Society for the Promotion of Science (Wakate B), the Kyoto University Hakubi Project, and the Research Center Network for Realization of Regenerative Medicine, Program for Intractable Diseases Research Utilizing Disease-Specific iPS Cells, of the Japan Science and Technology Agency (JST).

References

1. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. Embryonic stem cell lines derived from human blastocysts. Science. 1998;282(5391):1145–7. Epub 1998/11/06.
2. Gertow K, Przyborski S, Loring JF, Auerbach JM, Epifano O, Otonkoski T, et al. Isolation of human embryonic stem cell-derived teratomas for the assessment of pluripotency. Curr Protoc Stem Cell Biol. 2007;Chap. 1:Unit1B 4. Epub 2008/09/12.
3. Keller G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. Genes Dev. 2005;19(10):1129–55. Epub 2005/05/21.
4. Outten JT, Gadue P, French DL, Diamond SL. High-throughput screening assay for embryoid body differentiation of human embryonic stem cells. Curr Protoc Stem Cell Biol. 2012;Chap. 1:Unit 1D 6. Epub 2012/03/15.
5. Zacharias DG, Nelson TJ, Mueller PS, Hook CC. The science and ethics of induced pluripotency: what will become of embryonic stem cells? Mayo Clin Proc. 2011;86(7):634–40. Epub 2011/07/02.
6. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663–76. Epub 2006/08/15.
7. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007;131(5):861–72. Epub 2007/11/24.
8. Takahashi K, Yamanaka S. A decade of transcription factor-mediated reprogramming to pluripotency. Nat Rev Mol Cell Biol. 2016;17(3):183–93. Epub 2016/02/18.
9. Inoue H, Nagata N, Kurokawa H, Yamanaka S. iPS cells: a game changer for future medicine. EMBO J. 2014;33(5):409–17. Epub 2014/02/07.
10. Hindley C, Philpott A. The cell cycle and pluripotency. Biochem J. 2013;451(2):135–43. Epub 2013/03/29.
11. Watanabe K, Ueno M, Kamiya D, Nishiyama A, Matsumura M, Wataya T, et al. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat Biotechnol. 2007;25(6):681–6. Epub 2007/05/29.

12. Nakagawa M, Taniguchi Y, Senda S, Takizawa N, Ichisaka T, Asano K, et al. A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. Sci Rep. 2014;4:3594. Epub 2014/01/09.
13. Chen G, Gulbranson DR, Hou Z, Bolin JM, Ruotti V, Probasco MD, et al. Chemically defined conditions for human iPSC derivation and culture. Nat Methods. 2011;8(5):424–9. Epub 2011/04/12.
14. Kim SI, Oceguera-Yanez F, Sakurai C, Nakagawa M, Yamanaka S, Woltjen K. Inducible transgene expression in human iPS cells using versatile all-in-one piggyBac transposons. Methods Mol Biol. 2016;1357:111–31. Epub 2015/05/31.
15. Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell. 1987;51(3):503–12. Epub 1987/11/06.
16. Capecchi MR. Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet. 2005;6(6):507–12. Epub 2005/06/03.
17. Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. Nat Biotechnol. 2003;21(3):319–21. Epub 2003/02/11.
18. Lieber MR, Lu H, Gu J, Schwarz K. Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. Cell Res. 2008;18(1):125–33. Epub 2007/12/19.
19. Jasin M. Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet. 1996;12(6):224–8. Epub 1996/06/01.
20. Carroll D. Using nucleases to stimulate homologous recombination. Methods Mol Biol. 2004;262:195–207. Epub 2004/02/11.
21. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A. 1996;93(3):1156–60. Epub 1996/02/06.
22. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 2010;11(9):636–46. Epub 2010/08/19.
23. Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res. 2000;28(17):3361–9. Epub 2000/08/23.
24. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science. 1991;252(5007):809–17. Epub 1991/05/10.
25. Maeder ML, Thibodeau-Beganny S, Sander JD, Voytas DF, Joung JK. Oligomerized pool engineering (OPEN): an 'open-source' protocol for making customized zinc-finger arrays. Nat Protoc. 2009;4(10):1471–501. Epub 2009/10/03.
26. Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. Curr Opin Biotechnol. 2012;23(5):644–50. Epub 2012/02/22.
27. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011;39(12), e82. Epub 2011/04/16.
28. Bogdanove AJ, Schornack S, Lahaye T. TAL effectors: finding plant genes for disease and defense. Curr Opin Plant Biol. 2010;13(4):394–401. Epub 2010/06/24.
29. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009;326(5959):1509–12. Epub 2009/11/26.
30. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009;326(5959):1501. Epub 2009/11/26.
31. Chen S, Oikonomou G, Chiu CN, Niles BJ, Liu J, Lee DA, et al. A large-scale in vivo analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. Nucleic Acids Res. 2013;41(4):2769–78. Epub 2013/01/11.
32. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315(5819):1709–12. Epub 2007/03/24.

33. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096): 816–21. Epub 2012/06/30.
34. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32(4):347–55. Epub 2014/03/04.
35. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471(7340):602–7. Epub 2011/04/02.
36. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823–6. Epub 2013/01/05.
37. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819–23. Epub 2013/01/05.
38. Niu J, Zhang B, Chen H. Applications of TALENs and CRISPR/Cas9 in human cells and their potentials for gene therapy. Mol Biotechnol. 2014;56(8):681–8. Epub 2014/05/30.
39. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014;343(6166):80–4. Epub 2013/12/18.
40. Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol. 2014;32(3):267–73. Epub 2014/02/19.
41. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. Nature. 2015;520(7546):186–91. Epub 2015/04/02.
42. Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. Proc Natl Acad Sci U S A. 2013;110(39):15644–9. Epub 2013/08/14.
43. Kleinstiver BP, Prew MS, Tsai SQ, Nguyen NT, Topkar VV, Zheng Z, et al. Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. Nat Biotechnol. 2015;33(12):1293–8. Epub 2015/11/03.
44. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015;163(3):759–71. Epub 2015/10/01.
45. Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, Vakoc CR. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat Biotechnol. 2015;33(6):661–7. Epub 2015/05/12.
46. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. 2014;24(1): 132–41. Epub 2013/11/21.
47. Oceguera-Yanez F, Kim SI, Matsumoto T, Tan GW, Xiang L, Hatani T, et al. Engineering the AAVS1 locus for consistent and scalable transgene expression in human iPSCs and their differentiated derivatives. Methods. 2015. Epub 2015/12/29.
48. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell. 2013;154(6):1370–9. Epub 2013/09/03.
49. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013;153(4):910–8. Epub 2013/05/07.
50. Papapetrou EP, Lee G, Malani N, Setty M, Riviere I, Tirunagari LM, et al. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nat Biotechnol. 2011;29(1):73–8. Epub 2010/12/15.
51. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol. 2009;27(9):851–7. Epub 2009/08/15.
52. Smith JR, Maguire S, Davis LA, Alexander M, Yang F, Chandran S, et al. Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. Stem Cells. 2008;26(2):496–504. Epub 2007/11/21.

53. Soriano P. Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet. 1999;21(1):70–1. Epub 1999/01/23.
54. Zou J, Sweeney CL, Chou BK, Choi U, Pan J, Wang H, et al. Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. Blood. 2011;117(21):5561–72. Epub 2011/03/18.
55. Chang CJ, Bouhassira EE. Zinc-finger nuclease-mediated correction of alpha-thalassemia in iPS cells. Blood. 2012;120(19):3906–14. Epub 2012/09/25.
56. Wang Y, Liang P, Lan F, Wu H, Lisowski L, Gu M, et al. Genome editing of isogenic human induced pluripotent stem cells recapitulates long QT phenotype for drug testing. J Am Coll Cardiol. 2014;64(5):451–9. Epub 2014/08/02.
57. Sakurai K, Shimoji M, Tahimic CG, Aiba K, Kawase E, Hasegawa K, et al. Efficient integration of transgenes into a defined locus in human embryonic stem cells. Nucleic Acids Res. 2010;38(7), e96. Epub 2010/01/15.
58. Inai Y, Ohta Y, Nishikimi M. The whole structure of the human nonfunctional L-gulono-gamma-lactone oxidase gene—the gene responsible for scurvy—and the evolution of repetitive sequences thereon. J Nutr Sci Vitaminol. 2003;49(5):315–9. Epub 2004/01/02.
59. Li X, Burnight ER, Cooney AL, Malani N, Brady T, Sander JD, et al. piggyBac transposase tools for genome engineering. Proc Natl Acad Sci U S A. 2013;110(25):E2279–87. Epub 2013/06/01.
60. Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat Methods. 2011;8(10):861–9. Epub 2011/08/23.
61. Cerbini T, Funahashi R, Luo Y, Liu C, Park K, Rao M, et al. Transcription activator-like effector nuclease (TALEN)-mediated CLYBL targeting enables enhanced transgene expression and one-step generation of dual reporter human induced pluripotent stem cell (iPSC) and neural stem cell (NSC) lines. PLoS One. 2015;10(1), e0116032. Epub 2015/01/15.
62. Sharma R, Anguela XM, Doyon Y, Wechsler T, DeKelver RC, Sproul S, et al. In vivo genome editing of the albumin locus as a platform for protein replacement therapy. Blood. 2015;126(15):1777–84. Epub 2015/08/25.
63. Kim HS, Bernitz JM, Lee DF, Lemischka IR. Genomic editing tools to model human diseases with isogenic pluripotent stem cells. Stem Cells Dev. 2014;23(22):2673–86. Epub 2014/07/31.
64. Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. Blood. 2011;118(17):4599–608. Epub 2011/09/02.
65. Meier ID, Bernreuther C, Tilling T, Neidhardt J, Wong YW, Schulze C, et al. Short DNA sequences inserted for gene targeting can accidentally interfere with off-target gene expression. FASEB J. 2010;24(6):1714–24. Epub 2010/01/30.
66. Davis RP, Costa M, Grandela C, Holland AM, Hatzistavrou T, Micallef SJ, et al. A protocol for removal of antibiotic resistance cassettes from human embryonic stem cells genetically modified by homologous recombination or transgenesis. Nat Protoc. 2008;3(10):1550–8. Epub 2008/09/20.
67. Grindley ND, Whiteson KL, Rice PA. Mechanisms of site-specific recombination. Annu Rev Biochem. 2006;75:567–605. Epub 2006/06/08.
68. Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hamalainen R, et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature. 2009; 458(7239):766–70. Epub 2009/03/03.
69. Yusa K, Rashid ST, Strick-Marchand H, Varela I, Liu PQ, Paschon DE, et al. Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells. Nature. 2011;478(7369):391–4. Epub 2011/10/14.
70. Kim H, Kim K, Kim J, Kim SH, Yim J. Mutagenesis by imprecise excision of the piggyBac transposon in *Drosophila melanogaster*. Biochem Biophys Res Commun. 2012;417(1):335–9. Epub 2011/12/14.

71. Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res. 2010;38(15), e152. Epub 2010/06/10.
72. Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nat Methods. 2011;8(9):753–5. Epub 2011/07/19.
73. Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, et al. A TALEN genome-editing system for generating human stem cell-based disease models. Cell Stem Cell. 2013;12(2):238–51. Epub 2012/12/19.
74. Soldner F, Laganiere J, Cheng AW, Hockemeyer D, Gao Q, Alagappan R, et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. Cell. 2011;146(2):318–31. Epub 2011/07/16.
75. Radecke S, Radecke F, Cathomen T, Schwarz K. Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. Mol Ther. 2010;18(4):743–53. Epub 2010/01/14.
76. Howden SE, Maufort JP, Duffin BM, Elefanty AG, Stanley EG, Thomson JA. Simultaneous reprogramming and gene correction of patient fibroblasts. Stem Cell Rep. 2015;5(6):1109–18. Epub 2015/11/21.
77. Miyaoka Y, Chan AH, Judge LM, Yoo J, Huang M, Nguyen TD, et al. Isolation of single-base genome-edited human iPS cells without antibiotic selection. Nat Methods. 2014;11(3):291–3. Epub 2014/02/11.
78. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol. 2016;34:339–44. Epub 2016/01/21.
79. Smith C, Gore A, Yan W, Abalde-Atristain L, Li Z, He C, et al. Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. Cell Stem Cell. 2014;15(1):12–3. Epub 2014/07/06.
80. Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, et al. Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell. 2014;15(1):27–30. Epub 2014/07/06.
81. Kim HJ, Lee HJ, Kim H, Cho SW, Kim J-S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res. 2009;19:1279–88.
82. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. 2014;42(22), e168. Epub 2014/10/11.
83. Fujita T, Fujii H. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. Biochem Biophys Res Commun. 2013;439(1):132–6. Epub 2013/08/15.
84. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol. 2014;32(7):670–6. Epub 2014/04/23.
85. Frock RL, Hu J, Meyers RM, Ho YJ, Kii E, Alt FW. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat Biotechnol. 2015;33(2):179–86. Epub 2014/12/17.
86. Gabriel R, Lombardo A, Arens A, Miller JC, Genovese P, Kaeppel C, et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol. 2011;29(9):816–23. Epub 2011/08/09.
87. Wang X, Wang Y, Wu X, Wang J, Wang Y, Qiu Z, et al. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. Nat Biotechnol. 2015;33(2):175–8. Epub 2015/01/20.
88. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol. 2015;33(2):187–97. Epub 2014/12/17.

89. Crosetto N, Mitra A, Silva MJ, Bienko M, Dojer N, Wang Q, et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. Nat Methods. 2013;10(4):361–5. Epub 2013/03/19.
90. Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat Methods 2015;12(3):237–243, 244 Epub 2015/02/11.
91. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol. 2013; 31(9):822–6. Epub 2013/06/25.
92. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 2014;24(6):1012–9. Epub 2014/04/04.
93. Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, et al. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol. 2015;208:44–53. Epub 2015/05/25.
94. Merkle FT, Neuhausser WM, Santos D, Valen E, Gagnon JA, Maas K, et al. Efficient CRISPR-Cas9-mediated generation of knockin human pluripotent stem cells lacking undesired mutations at the targeted locus. Cell Rep. 2015;11(6):875–83. Epub 2015/05/06.
95. Sanders KL, Catto LE, Bellamy SR, Halford SE. Targeting individual subunits of the FokI restriction endonuclease to specific DNA strands. Nucleic Acids Res. 2009;37(7):2105–15. Epub 2009/02/19.
96. Ramirez CL, Certo MT, Mussolino C, Goodwin MJ, Cradick TJ, McCaffrey AP, et al. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. Nucleic Acids Res. 2012;40(12):5560–8. Epub 2012/03/01.
97. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013;154(6):1380–9. Epub 2013/09/03.
98. Certo MT, Ryu BY, Annis JE, Garibov M, Jarjour J, Rawlings DJ, et al. Tracking genome engineering outcome at individual DNA breakpoints. Nat Methods. 2011;8(8):671–6. Epub 2011/07/12.
99. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014;156(5):935–49. Epub 2014/02/18.
100. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014;343(6176):1247997. Epub 2014/02/08.
101. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature. 2016;529(7587):490–5. Epub 2016/01/07.
102. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol. 2014;32(3):279–84. Epub 2014/01/28.
103. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol. 2014;32(12):1262–7. Epub 2014/09/04.
104. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016;34(2):184–91. Epub 2016/01/19.
105. Murry CE, Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell. 2008;132(4):661–80. Epub 2008/02/26.
106. Merkle FT, Eggan K. Modeling human disease with pluripotent stem cells: from genome association to function. Cell Stem Cell. 2013;12(6):656–68. Epub 2013/06/12.
107. Kondo T, Asai M, Tsukita K, Kutoku Y, Ohsawa Y, Sunada Y, et al. Modeling Alzheimer's disease with iPSCs reveals stress phenotypes associated with intracellular Aβ and differential drug responsiveness. Cell Stem Cell. 2013;12(4):487–96. Epub 2013/02/26.

108. Zeltner N, Studer L. Pluripotent stem cell-based disease modeling: current hurdles and future promise. Curr Opin Cell Biol. 2015;37:102–10. Epub 2015/12/03.
109. Osafune K, Caron L, Borowiak M, Martinez RJ, Fitz-Gerald CS, Sato Y, et al. Marked differences in differentiation propensity among human embryonic stem cell lines. Nat Biotechnol. 2008;26(3):313–5. Epub 2008/02/19.
110. Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, et al. Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell. 2011;144(3):439–52. Epub 2011/02/08.
111. Ronen D, Benvenisty N. Genomic stability in reprogramming. Curr Opin Genet Dev. 2012;22(5):444–9. Epub 2012/10/09.
112. Bhutani K, Nazor KL, Williams R, Tran H, Dai H, Dzakula Z, et al. Whole-genome mutational burden analysis of three pluripotency induction methods. Nat Commun. 2016;7:10536. Epub 2016/02/20.
113. Abyzov A, Mariani J, Palejev D, Zhang Y, Haney MS, Tomasini L, et al. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature. 2012;492(7429):438–42. Epub 2012/11/20.
114. Young MA, Larson DE, Sun CW, George DR, Ding L, Miller CA, et al. Background mutations in parental cells account for most of the genetic heterogeneity of induced pluripotent stem cells. Cell Stem Cell. 2012;10(5):570–82. Epub 2012/05/01.
115. Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, et al. Epigenetic memory in induced pluripotent stem cells. Nature. 2010;467(7313):285–90. Epub 2010/07/21.
116. Nasu A, Ikeya M, Yamamoto T, Watanabe A, Jin Y, Matsumoto Y, et al. Genetically matched human iPS cells reveal that propensity for cartilage and bone differentiation differs with clones, not cell type of origin. PLoS One. 2013;8(1), e53771. Epub 2013/02/06.
117. Kyttala A, Moraghebi R, Valensisi C, Kettunen J, Andrus C, Pasumarthy KK, et al. Genetic variability overrides the impact of parental cell type and determines iPSC differentiation potential. Stem Cell Rep. 2016;6(2):200–12. Epub 2016/01/19.
118. Turner M, Leslie S, Martin NG, Peschanski M, Rao M, Taylor CJ, et al. Toward the development of a global induced pluripotent stem cell library. Cell Stem Cell. 2013;13(4):382–4. Epub 2013/10/08.
119. Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M. Genome-wide association studies in diverse populations. Nat Rev Genet. 2010;11(5):356–66. Epub 2010/04/17.
120. Yusa K. Seamless genome editing in human pluripotent stem cells using custom endonuclease-based gene targeting and the piggyBac transposon. Nat Protoc. 2013;8(10):2061–78. Epub 2013/09/28.
121. Woltjen K, Hamalainen R, Kibschull M, Mileikovsky M, Nagy A. Transgene-free production of pluripotent stem cells using piggyBac transposons. Methods Mol Biol. 2011;767:87–103. Epub 2011/08/09.

CRISPR/Cas9 and the Paradigm Shift in Mouse Genome Manipulation Technologies

Channabasavaiah B. Gurumurthy, Rolen M. Quadros, Masahiro Sato, Tomoji Mashimo, K.C. Kent Lloyd, and Masato Ohtsuka

Abstract The CRISPR revolution that began in 2013 has been adopted and embraced by many researchers worldwide, including the mouse molecular genetics community. CRISPR represents one of only a few radical and transformative shifts in transgenic technologies over the past 30 years. This chapter discusses the paradigm shift that CRISPR technology has brought about in the field of mouse genome editing.

Keywords Transgenic • Knockout • Knock-in • Genome editing • CRISPR/Cas9 • Mouse genome • Microinjection

C.B. Gurumurthy (✉)
Developmental Neuroscience, Munroe Meyer Institute for Genetics and Rehabilitation, University of Nebraska Medical Center, Omaha, NE, USA

Mouse Genome Engineering Core Facility, Office of the Vice Chancellor for Research, University of Nebraska Medical Center, Omaha, NE, USA
e-mail: cgurumurthy@unmc.edu

R.M. Quadros
Mouse Genome Engineering Core Facility, Office of the Vice Chancellor for Research, University of Nebraska Medical Center, Omaha, NE, USA

M. Sato
Section of Gene Expression Regulation, Frontier Science Research Center, Kagoshima University, Kagoshima 890-8544, Japan

T. Mashimo
Institute of Experimental Animal Sciences, Graduate School of Medicine, Osaka University, Osaka 565-0871, Japan

K.C.K. Lloyd
Mouse Biology Program, University of California, Davis, 2795 2nd Street, Davis, CA 95618, USA

M. Ohtsuka (✉)
Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, Tokai University School of Medicine, 143 Shimokasuya, Isehara, Kanagawa 259-1193, Japan
e-mail: masato@is.icc.u-tokai.ac.jp

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_4

Introduction

During the past three decades, techniques and procedures in genome manipulation developed and evolved primarily using the laboratory mouse as a model system, mainly because of the availability of murine embryonic stem (ES) cells. ES cells from no species other than the mouse were as robust and efficient for usurping homologous recombination (HR) to induce targeted genetic changes in the mammalian genome. Methods for targeted genomic manipulation without the use of mouse ES cells were practically nonexistent. In the last few years, many techniques involving "designer nucleases" such as zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats/Crispr-associated 9 (CRISPR/Cas9) have been developed that enable species-agnostic genome editing without the need for ES cells. These new techniques have breached the species barrier and seamlessly made their way into the genome-editing arenas of many other species.

Among the methods that use designer nucleases (also known as sequence-specific endonucleases), the CRISPR/Cas9 system has become the most popular. Its technical simplicity, rapidity of designing and performing experiments, and high success rate have been well documented in almost every species in which it has been tried to date. In the pre-CRISPR era, ES cells served as critical reagents enabling giant strides in the field, including the development of tens of thousands of reagents to systematically knock out (KO) mouse genes. Now, however, their utility is becoming overshadowed by new technologies as the mouse genome engineering community is shifting heavily toward CRISPR-based genome-editing approaches. In this chapter, we discuss how this novel technology has impacted the field.

Traditional Mouse Genome Engineering Technologies in the Pre-CRISPR Era

The development of mouse genetic engineering began in the 1980s by attempting to transfer exogenous DNA (genes) into a genome for developing "transgenic (Tg)" animals and to delete/inactivate endogenous genes for developing "knockout (KO)" animals. Transgenesis was achieved by microinjecting purified Tg DNA (transgenes) into fertilized zygotes and subsequently by transferring the zygotes to pseudo-pregnant recipient animals to generate live animals. To accomplish gene knockouts, however, simple injection of DNA into zygotes (as done in the case of transgenics) would not be enough: it was necessary to develop a special tool, the ES cells. Using ES cells, the endogenous genes were first modified through the HR process (which occurs during DNA repair in cells); in the second step the ES cells containing the modified gene were injected into blastocyst embryos about 3 days old to generate chimeric animals [1]. The resulting chimeric animals would contain cells originating from two sources: the cells derived from the embryo host blastocyst and those from exogenously injected ES cells. Breeding of the chimera to a wild-type mouse would result in vertical transmission of the gene-modified allele (the ES cell-derived mutant allele) to heterozygous offspring. Intercrossing between heterozygous offspring would result in

production of homozygous, fully ES cell-derived mice. The third type of genetically engineered animal model is a *Knockin* (knock-in, KI), which refers to (usually) targeted insertion of DNA with a desired type of genetic change. Generation of KI mice also required ES cells, and follows similar steps as just outlined.

The traditional Tg and KO/KI techniques prevailed for nearly three decades and have helped generate thousands of mouse disease models. However, these time-tested methods have a few limitations that are discussed next (sections "Traditional Tg Techniques and Their Limitations" to "Microinjection and Its Limitations").

Traditional Tg Techniques and Their Limitations

Traditional Tg mice are generated by direct microinjection of Tg DNA (that consists of elements such as promoter, cDNA, transcription terminator, etc.) into the pronuclei of 0.5-day-old fertilized eggs (also known as zygotes), followed by their subsequent transfer to oviducts of pseudo-pregnant mice. The live offspring obtained are called founder (G_0) animals if they contain the DNA of interest, which are then bred to establish the Tg line. The detailed descriptions of designing and generating KO/KI mice are reviewed by Haruyama and Kulkarni (Haruyama et al. [2]).

The traditional Tg mice generation methods have a few inherent limitations: (1) random integration of the transgene where local regulatory elements could affect its expression, and/or the transgene itself can disrupt or affect the expression of local genes, and (2) integration at multiple sites or multiple copy integration, which occasionally result in unreliable expression or transgene silencing [3]. Because of such pitfalls, several Tg G_0 lines are screened for desired expression before the lines are established for further experiments, a tedious but necessary step using random integration-based Tg mice generation projects [4].

Traditional KO/KI Techniques and Their Limitations

The traditional KO/KI models were generated through the use of ES cells that allow HR to replace or insert a genetically engineered DNA copy of a recombinant DNA construct that is designed and built for each KO/KI project. The process, in brief, includes four major steps: (1) construction of molecular targeting construct; (2) electroporation of the targeting construct into ES cells followed by positive/negative selection of correctly targeted clones; (3) microinjection of ES cell clones into blastocysts and transfer into pseudo-pregnant recipients to generate chimeras; and (4) breeding of chimeras with wild-type mouse to obtain a germline-transmitted mouse line. Detailed descriptions of designing and generating KO/KI mice are reviewed by Hall et al. [5].

Traditional KO/KI mice generation methods also have a few inherent limitations. First, ES cells must retain pluripotency to populate germ cells and vertically transmit the induced mutation to the next generation. Second, germline-competent ES cells are available for only a very few genetic backgrounds, not for many of the commonly used mouse strains. Animal models that cannot be generated in a pure genetic background for many of those strains must undergo many generations of backcross breeding to achieve

congenesis [6]. Third, efficient insertion of the targeting construct requires long regions of DNA, or homology arms, flanking each end of the intended target, which can be difficult to achieve for certain genes. Fourth, the targeting constructs need to contain additional elements, such as positive selection (e.g., neomycin or puromycin resistance genes) and negative selection (e.g., thymidine kinase or diphtheria toxin) markers to select single-cell clones that contain the correctly inserted DNA. Fifth, the insertion of long positive selection markers [or certain genetic elements such as flippase (Flp) or Cre recombinase sites that flank the positive selection cassettes] may result in unintentional interference of the regulatory elements near the gene locus. Sixth, the KO/KI strategy would be difficult if a conditional KO needs to be developed for single-exon genes. Seventh, not all chimeras result in germ line-transmitted offspring. Eighth, traditional gene targeting can generally only be used to generate no more than one gene KO/KI in an experiment. Last, design and generation of KO/KI animal models is labor intensive, requires extensive amounts of time, and is quite expensive.

Microinjection and Its Limitations

Both Tg and KO/KI techniques require microinjection directly into mouse embryos. Although tedious, labor intensive, and expensive, the microinjection technique has been used as the gold standard for more than three decades for developing genetically engineered mouse models. The desired DNA cassette is microinjected into zygotes for generation of Tg mice, whereas gene-targeted ES cells are microinjected into blastocysts for generation of KO/KI mice. Zygotes or embryos are produced from females that are superovulated and mated with stud males. To ensure a sufficient number of Tg G_0 lines or chimeras, typically 100 or more eggs or 50 or more embryos are injected for Tg or KO/KI projects, respectively. The manipulated embryos need to be surgically transferred into pseudo-pregnant females to generate live offspring.

In general, microinjection has been an integral step in mouse gene targeting projects, but its two major limitations are that it requires sophisticated equipment and well-trained and experienced personnel to perform the procedure. Typically, microinjection equipment costs about $100,000–$200,000, and microinjection (and associated embryo-handling techniques) requires significant practice to perfect. At least a couple years of regular practice are required for a researcher to learn and be proficient in performing microinjection. Also, one needs continued practice to retain technical proficiency.

CRISPR/Cas9 and Mouse Genome Editing

Since 2013, the CRISPR-mediated genome editing has revolutionized almost every field of biology. Briefly, it uses a single guide RNA (sgRNA) with a 20-nucleotide sequence complementary to the target site in the genome, to bring a Cas9 nuclease to the site and make a double-stranded DNA break. This break is then repaired through an error-prone cellular DNA repair process called nonhomologous end-joining (NHEJ), resulting in gene disruption. The cut site can also be repaired by

providing a short DNA oligonucleotide with homology regions of about 30–60 nucleotides, or by providing a repair template with long homology arms (typically about 0.5–2 kb or longer). Such donor DNAs are inserted through the less efficient repair processes, such as homology-directed repair (HDR) [7] for single-stranded short templates or HR when using double-stranded templates.

CRISPR technology can also be regarded as a "disruptor" because it has changed the basic format of how Tg and animal genome engineering experiments are performed, which has essentially remained unchanged during the past three decades. Additionally, CRISPR has superseded the other two genome-editing methods, ZFNs and TALENs, which prevailed for about 3 to 5 years before the CRISPR era. Because of its simplicity, relative ease, and rapidity to manipulate the genome, the CRISPR/Cas9 technique has propelled many technology developers to think outside the box and devise novel and innovative features that make it versatile and adaptable to diverse fields of research. As we usher in a new era of genome engineering driven by CRISPR-related techniques, section "CRISPR Technology and the Paradigm Shifts in Mouse Genome Engineering" discusses the paradigm shifts in the field.

Box 1: Limitations of Traditional Mouse Genetic Engineering Technologies and the Paradigm Shifts Created by CRISPR/Cas9 Genome Editing

Traditional genetic engineering approaches	The CRISPR/Cas9 genome-editing approach
ES cells are absolutely essential for the generation of KO/KI models (costly and time consuming)	Can generate KO/KI models without the need for ES cells (cost- and time saving)
Techniques can be limited to certain strains where ES cells are available, particularly for KO/KI models	Can develop KO/KI models under any strain background
Difficult to generate KO/KI models without inserting additional elements in the genome	Can generate most KO/KI models without inserting additional elements in the genome
Generation of homozygous KO/KI mutants in G_0 stage is not possible and G_0 animals need to be intercrossed to obtain homozygotes before they are used for phenotyping	Generation of homozygous KO/KI mutants in G_0 stage is readily possible, and they can be used for direct phenotyping in some cases (for instance, any visible traits, hematological phenotypes)
Except in case of advanced techniques [8, 9], multiplexing (multiple Tg lines or multiple genes KO/KI) is difficult	Multiplexing is readily possible
Except in case of advanced techniques [8, 9], pronuclear injection of Tg DNA will get integrated randomly, often more than one copy and/or occasionally at multiple locations	Targeted insertion of single copy at Cas9 cut site is possible
Large-scale genome modifications (deletions or replacements) are difficult	Large-scale genome modifications are readily possible
Microinjection is a critical step; each embryo must be injected manually and transferred back into recipient females	Electroporation can replace the microinjection step and many embryos can be processed simultaneously [10–12]. More advanced approaches (GONAD) can even obviate the need for ex vivo embryo handling [13]

CRISPR Technology and the Paradigm Shifts in Mouse Genome Engineering

Clearly, the CRISPR system has impacted traditional Tg and KO technologies. Many Tg mouse labs and core facilities across the globe have added CRISPR to their toolbox. The paradigm shifts in mouse genome editing that are listed in Box 1 are discussed next.

1. *Ability to bypass the use of ES cells*. Two important features of ES cells that make them critical reagents for genome engineering are (1) they maintain pluripotency during culturing and gene targeting and (2) they enable high competency in HR. Historically, mouse ES were the only ES cells with these two features that showed robust performance. Attempts to establish ES cells for other species (except rats) have failed to date. One of the biggest paradigm shifts that CRISPR has caused in the field is its ability to bypass the need for ES cells. This, along with simplicity and lower cost, is the main reason why CRISPR has been so widely applicable in creating gene KO models in any species. Even in mice, with the advent of CRISPR, ES cells that served as valuable tools in generating thousands of mouse models during the past two to three decades are now being superseded by the use of CRISPR [14].
2. *Ability to generate KO/KI mice on any genetic background*. Because CRISPR-mediated gene editing can be used directly on zygotes to edit genes, practically any strain of mouse can be used for generating KO/KI models. Previously, the field relied on the availability of strain-specific ES cells for developing mouse models. Although better-quality ES cells for the C57BL/6N strain (the most popular in disease research) were developed during the past decade [15], for many years mouse KO technology predominantly relied on ES cells derived from sub-strains on the 129 genetic background. G_0 lines generated using 129 ES cells injected into a different genetic background (e.g., C57BL/6) are a mixed strain background that required backcrossing to the desired genetic background for many generations before the model could be used for experiments. The CRISPR system readily offers solutions to such limitations, as it is applicable to any strain, thus obviating the need for backcrossing.
3. *Ability to generate point mutations without any other genetic disruptions*. For generation of simple KI models such as creating point mutations to mimic human disease or restoring the function of mutant proteins, the CRISPR/Cas9 system offers distinct advantages over traditional methods. Specifically, point mutations can be inserted without the need to include extra DNA near the locus, such as a positive selection cassette when using ES cell-based methods. Occasionally, the presence of such extra elements near the locus may affect gene expression by disrupting adjacent yet unknown regulatory elements, etc.
4. *Ability to generate homozygous mutant mice in F_0 generation*. With traditional approaches using either ES cells or random transgenesis, it was not possible to obtain homozygous G_0 animals. CRISPR-generated models can produce homozygous mutations in the G_0 generation and can be used for a quick phenotypic analysis, albeit mosacisim and possible off-target effects must be considered as

confounders. Nevertheless, in some cases, phenotypic screening of G_0 progeny can provide significant cost and time savings, when compared to the use of other methods that require breeding to achieve homozygosity.

5. *Ability to generate multiple mutations in one microinjection experiment*. Despite their popular use in the pre-CRISPR era, traditional KO/KI approaches are inadequate in the following aspects: (1) it is practically impossible to simultaneously generate KO mutations for more than one gene at a time; (2) germline transmission of the mutant allele is often not guaranteed; and (3) the models would be only heterozygous initially. In comparison, CRISPR-mediated gene editing offers giant solutions to these limitations. Indeed, the generation of up to five KO mutant models has been reported in one session [16] with CRISPR, where previously such a task would take more than 3 to 4 years because of the time-consuming breeding steps after generating individual KO mice. The cost for such traditional KO projects would be severalfold more than that of CRISPR-based approaches because of the lengthy steps involved. Using CRISPR, it is not uncommon to obtain homozygous alleles for some mutations.
6. *Large-scale genome modifications*. Although mouse models of large chromosomal deletions and insertions of hundreds of kilobases have been developed using traditional ES cell-based approaches [17, 18], clearly such projects need enormous amounts of resources and time to accomplish, because they were performed through a series of complex and successive modifications. Using certain advanced CRISPR-based strategies, such large-scale insertions and deletions are now possible, making the system highly cost effective [19].
7. *Cytoplasmic microinjection*. The traditional Tg models are developed by injecting Tg DNA into pronuclei because the injected DNA is intended to be inserted into the genome. Because the CRISPR system constitutes a sgRNA and a Cas9 endonuclease, pronuclear injection might not be required (which can be a difficult skill to master). Further, pronuclei in certain strains of mice are not easy to visualize for microinjection. In many cases, cytoplasmic injection seems to be sufficient [20], especially in cases where simple indel mutants are to be generated without the need for coinjection of complex donor DNA templates. When combined with a donor DNA template for genomic insertion, it is necessary to deliver injection mixture to the pronucleus to ensure insertion efficiency. Simultaneous cytoplasmic and pronuclear injection has become a popular strategy in many labs for CRISPR-based genome-editing applications that suit both NHEJ and HR mechanisms.
8. *Novel delivery approaches*. CRISPR tools can be delivered to embryos without the need for microinjection or ex vivo handling of embryos. Although microinjection has been used as the gold standard for more than three decades for developing genetically engineered mouse models, there has been constant effort by many researchers to develop microinjection-independent methods because of the inherent limitations of microinjection (covered in the section "Microinjection and Its Limitations"). The advent of CRISPR readily enabled the development of an electroporation technique that can be performed on several embryos at a time, instead of manually injecting them one by one [10–12]. A step further is a new technique developed by us called Genome Editing via Oviductal Nucleic Acids

Delivery (GONAD) [13]. GONAD allows direct electroporation of CRISPR tools into embryos in situ without the need for embryo isolation and handling ex vivo. Thus, GONAD serves to bypass all major bottlenecks of animal transgenesis: isolation, microinjection, and surgical transfer of embryos into pseudopregnant mice [13, 21, 22].

9. *Germline transmission potential can be high compared to traditional methods-derived chimeras*. Because the traditional ES cell-based approach relies on the pluripotency and germline transmission potential of ES cells, some chimeras may not result in passage of the targeted allele to the next generation of offspring. Reasons for failure of germline transmission include low contribution of ES-derived germ cells in chimeric mice, or loss of ES cell pluripotency. Although mosaicism remains a potential disadvantage, the germ cells of CRISPR-generated G_0 mice are expected to contain CRISPR-induced mutation(s). Therefore, the chances of germline transmission of a CRISPR-induced mutant allele to the next generation of offspring is high. Furthermore, certain CRISPR-generated G_0 mice may contain two or more types of mutations at a given locus, which can be segregated by breeding. Even though the segregation process seems complicated in certain cases, multiple different mutations at the given locus offer more options to study the phenotype using multiple alleles.

The Current Challenges of CRISPR/Cas9-Mediated Mouse Genome Engineering

Poor Efficiency of Insertion of Sequences at Cas9 Cut Sites

Although there are a few reports that demonstrate the insertion of longer DNA cassettes at Cas9-cut sites, increasing the overall efficiency of insertion is an area that needs further development. Despite its widespread use, to achieve targeted insertion at many loci still remains challenging. Although one of the problems may be less efficient guide sequences, the overall low insertion efficiency may also be attributed to the loci (e.g., extent of chromatin density) and donor DNA design (e.g., extent of genomic homology to target region). Additional strategies are necessary to make the CRISPR system suitable for efficient insertion of large DNA cassettes and for generating models more complex than indels on a routine basis.

Challenges in Developing Conditional KO Models

Conditional KO mouse models with two *loxP* sites flanking the target exon/s is a standard approach followed in traditional ES cell-based applications. Many labs have been trying to develop conditional KO models using CRISPR. Insertion of two

loxP sites can be achieved through one of two ways: (1) using a double-stranded (ds) DNA donor containing short homology arms (~1 kb) and two *loxP* sites flanking the target region [23], or (2) using two separate ssODNs (single-stranded oligodeoxynucleotides) encoding *loxP* sites in the middle and ultrashort (~60-base-pair) flanking homology arms corresponding to the desired genomic sites and inserting them through two separate CRISPR cuts in the genome [24]. High-efficiency insertion of two *loxP* sites in *cis* orientation remains elusive and challenging. The reasons for this are (1) the two independent gRNAs should be efficient in causing double-strand breaks at their target sites; if one fails to cut, the process will not result in the desired alleles; (2) even if both guides work efficiently, NHEJ is still favored, causing the two flanking ends to join together and excluding the intermediary piece of DNA; and (3) challenges in genotyping of correctly inserted *loxP* sites, specifically using the two ssODNs approach (see the section "Challenges Associated with Genotyping").

Off-Target Effects

Because the gRNA recognition sequence is only 20 nucleotides long and certain mismatches are tolerated when gRNA binds to genomic DNA, use of CRISPR can result in unintentional off-target cleavages. Some of the initial studies, done in cell culture systems, cautioned that off-target effects could be a major concern with the use of CRISPR technology [25, 26]. Certain strategies have been described to minimize or eliminate off-target cleavages. (1) The Cas9 nickase (nCas9 or Cas9n) approach [27, 28] that uses a mutated Cas9 which can create a nick instead of a double-strand break; by using paired nickases, two nicks are created using two gRNAs close to the target site. (2) Delivery of Cas9 in the form of mRNA or protein instead of plasmid; continued Cas9 expressed from plasmid DNA would result in an abundance of Cas9 protein over a much longer period than needed, resulting in the potential for more off-target cuts than when using Cas9 mRNA or protein. It was presumed that off-target cleavages would be high, based on the observations made in cell culture systems. However, some recent reports demonstrate that off-target effects are minimal or nil in mouse models generated through the CRISPR system [29]; one of the main reasons for this is the use of Cas9 mRNA or protein [30, 31]. Further, the concern about off-target cleavages in mouse models can be addressed by backcrossing G_0 mice to segregate mutations through successive breeding steps.

Challenges Associated with Genotyping

Genotyping of CRISPR-generated offspring is another major challenge because it can generate many unexpected outcomes such as imprecise insertion of the donor template, and co-occurrence of more than two types of alleles (also known as mosaicism). It may require careful analysis of many offspring generated from G_0 mice to

segregate and establish the desired mutations. Also, genotyping may not be easy in certain cases using a simple PCR assay and it may require sequencing of every offspring. To avoid such complications, one can consider choosing gRNAs close to restriction endonuclease (RE) recognition sites or to include an RE site in their donor template to aid in designing RFLP-PCR (restriction fragment length polymorphism)-based genotyping. Genotyping in case of the *loxP* ssODN approach is particularly challenging to ensure correct insertion of *loxP* sites on the same allele. Specifically, genotypic discrimination of correct targeting can be challenging if (a) the two *loxP* sites are far apart and it is difficult to amplify the entire floxed region by PCR (for confirmation by RFLP), or if (b) if the G_0 animals are not homozygous for at least one of the insertion sites. In such scenarios, Southern blotting becomes necessary for accurate confirmation of *loxP* insertions on the same allele.

Future Impact of CRISPR/Cas9 on Manipulating the Mouse Genome

Clearly, CRISPR/Cas9 has revolutionized many fields of biology, including mouse genome manipulation. Newer CRISPR tools and improved strategies are constantly being added. A few more CRISPR nucleases were recently discovered [32, 33] that offer additional features, refinements, and capabilities to the CRISPR genome-editing toolbox. Such improvements can have a significant impact on both traditional Tg and KO/KI technologies.

Impact on Random Tg Technologies

The majority of Tg mouse models generated to date are of random Tg type; in many cases, such projects fail to obtain reliable TgG_0 lines with high efficiency. If CRISPR/Cas9 can be further improved to efficiently insert larger DNA cassettes into the genome at safe harbor sites (e.g., *ROSA26*), it is very likely that the community will shift to "CRISPR transgenesis" and eventually random integration-based Tg mice production may become obsolete. Although there are not many reports of successful insertion of longer DNA cassettes with CRISPR/Cas9, certain strategies described recently promise targeted transgenesis of larger cassettes [31, 34–36].

Impact on KO/KI Technologies

As already noted, the mouse molecular genetics field has been transitioning rapidly to using CRISPR/Cas9 for making point mutation KIs. Although it is demonstrated that conditional KO models can be generated using the CRISPR system [24], it has not yet become a commonly used method because of the inherent difficulties

associated with inserting two *loxP* sites flanking the target site in *cis*. It is likely that, in the near future, technical advances will evolve to develop conditional KO models easily and efficiently.

Impact on Microinjection Technique

Last, if in vitro zygote electroporation and GONAD techniques become popular, CRISPR-based methods do not need the specialized microinjection setup or specially skilled personnel, which can allow many researchers to perform genome-editing experiments, in contrast to specialized core facilities that performed traditional, microinjection-based genome editing. GONAD is a promising new method for delivering CRISPR reagents directly to zygotes within the oviducts through electroporation. Compared to microinjection, GONAD requires a higher concentration of reagents to ensure embryonic uptake and activity. Its wide applicability and use in future is likely to result in faster and novel evolution and refining of the CRISPR technique itself, facilitating a transformation in the field of genome editing.

Conclusion

Traditional mouse genome manipulation techniques, established over the past three decades, have been used to develop thousands of mouse models. The recent addition of genome-editing tools such as ZFNs, TALENs, and CRISPR/Cas9 have resulted in a rapid transformation in the landscape of genome manipulation. In particular, CRISPR/Cas9 has been widely adopted during the past 2 to 3 years, and its simplicity and applicability across species has made the process faster, more cost effective, and versatile. It has also helped technology developers to devise newer methodologies that would have been practically impossible in the pre-CRISPR era. Research is underway to find additional CRISPR endonuclease molecules, and newer strategies to improve DNA insertion efficiency and to facilitate and improve the insertion of longer DNA sequences. All such improvements would enable this simple and ingenious method of gene editing to revolutionize biomedical research in the years to come.

Acknowledgments C.B.G. acknowledges funding support by the NIH grants: Institutional Development Award (IDeA) from the National Institutes for General Medical Sciences grant number P20GM103471 (NIGMS) and NIH Office of Research Infrastructure Programs (ORIP) grant number R24OD018546 (ORIP/DPCPSI). M.O. acknowledges the funding support by the 2014 Tokai University School of Medicine Research Aid and Grant-in-Aid for challenging Exploratory Research (15 K14371) from Japan Society for the Promotion of Science (JSPS).

References

1. Behringer R, Gertsenstein M, Nagy KV, Nagy A. Manipulating the mouse embryo: a laboratory manual. Cold Spring Harbor: CSHL Press; 2014.
2. Haruyama N, Cho A, Kulkarni AB. Overview: engineering transgenic constructs and mice. Curr Protoc Cell Biol. 2009;Chap. 19:Unit 19.10.
3. Garrick D, Fiering S, Martin DI, Whitelaw E. Repeat-induced gene silencing in mammals. Nat Genet. 1998;18:56–9.
4. Ohtsuka M, Miura H, Sato M, Kimura M, Inoko H, Gurumurthy CB. PITT: pronuclear injection-based targeted transgenesis, a reliable transgene expression method in mice. Exp Anim. 2012;61:489–502.
5. Hall B, Limaye A, Kulkarni AB. Overview: generation of gene knockout mice. Curr Protoc Cell Biol. 2009;Chap. 19:Unit 19.12 19.12.1–17.
6. Gurumurthy CB, Joshi PS, Kurz SG, Ohtsuka M, Quadros RM, Harms DW, Lloyd KCK. Validation of simple sequence length polymorphism regions of commonly used mouse strains for marker assisted speed congenics screening. Int J Genomics. 2015;2015:735845.
7. Peng Y, Clark KJ, Campbell JM, Panetta MR, Guo Y, Ekker SC. Making designer mutants in model organisms. Development. 2014;141:4042–54.
8. Ohtsuka M, Ogiwara S, Miura H, Mizutani A, Warita T, Sato M, Imai K, Hozumi K, Sato T, Tanaka M, Kimura M, Inoko H. Pronuclear injection-based mouse targeted transgenesis for reproducible and highly efficient transgene expression. Nucleic Acids Res. 2010;38, e198.
9. Ohtsuka M, Miura H, Mochida K, Hirose M, Hasegawa A, Ogura A, Mizutani R, Kimura M, Isotani A, Ikawa M, Sato M, Gurumurthy CB. One-step generation of multiple transgenic mouse lines using an improved Pronuclear Injection-based Targeted Transgenesis (i-PITT). BMC Genomics. 2015;16:274.
10. Kaneko T, Sakuma T, Yamamoto T, Mashimo T. Simple knockout by electroporation of engineered endonucleases into intact rat embryos. Sci Rep. 2014;4:6382.
11. Kaneko T, Mashimo T. Simple genome editing of rodent intact embryos by electroporation. PLoS One. 2015;10, e0142755.
12. Hashimoto M, Takemoto T. Electroporation enables the efficient mRNA delivery into the mouse zygotes and facilitates CRISPR/Cas9-based genome editing. Sci Rep. 2015;5:11315.
13. Takahashi G, Gurumurthy CB, Wada K, Miura H, Sato M, Ohtsuka M. GONAD: genome-editing via Oviductal Nucleic Acids Delivery system: a novel microinjection independent genome engineering method in mice. Sci Rep. 2015;5:11406.
14. Skarnes WC. Is mouse embryonic stem cell technology obsolete? Genome Biol. 2015;16:109.
15. Pettitt SJ, Liang Q, Rairdan XY, Moran JL, Prosser HM, Beier DR, Lloyd KC, Bradley A, Skarnes WC. Agouti C57BL/6N embryonic stem cells for mouse genetic resources. Nat Methods. 2009;6:493–5.
16. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013;153:910–8.
17. Yu Y, Bradley A. Mouse genomic technologies: engineering chromosomal rearrangements in mice. Nat Rev Genet. 2001;2:780–90.
18. Scheer N, Kapelyukh Y, Chatham L, Rode A, Buechel S, Wolf CR. Generation and characterization of novel cytochrome P450 Cyp2c gene cluster knockout and CYP2C9 humanized mouse lines. Mol Pharmacol. 2012;82:1022–9.
19. Yoshimi K, Kunihiro Y, Kaneko T, Nagahora H, Voigt B, Mashimo T. ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. Nat Commun. 2016;7:10431.
20. Horii T, Arai Y, Yamazaki M, Morita S, Kimura M, Itoh M, Abe Y, Hatada I. Validation of microinjection methods for generating knockout mice by CRISPR/Cas-mediated genome engineering. Sci Rep. 2014;4:4513.
21. Gurumurthy CB, Takahashi G, Wada K, Miura H, Sato M, Ohtsuka M. GONAD: a novel CRISPR/Cas9 genome editing method that does not require ex vivo handling of embryos. Curr Protoc Hum Genet. 2016;88:15.8.1–12.

22. Sato M, Ohtsuka M, Watanabe S, Gurumurthy CB. Nucleic acids delivery methods for genome editing in zygotes and embryos: the old, the new, and the old-new. Biol Direct. 2016;11(1):16.
23. Lee AY, Lloyd KCK. Conditional targeting of Ispd using paired Cas9 nickase and a single DNA template in mice. FEBS Open Bio. 2014;4:637–42.
24. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell. 2013;154:1370–9.
25. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol. 2013;31:822–6.
26. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol. 2013;31:839–43.
27. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013;154:1380–9.
28. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X, Skarnes WC. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Methods. 2014;11:399–402.
29. Iyer V, Shen B, Zhang W, Hodgkins A, Keane T, Huang X, Skarnes WC. Off-target mutations are rare in Cas9-modified mice. Nat Methods. 2015;12:479.
30. Harms DW, Quadros RM, Seruggia D, Ohtsuka M, Takahashi G, Montoliu L, Gurumurthy CB. Mouse genome editing using the CRISPR/Cas system. Curr Protoc Hum Genet. 2014;83:15.7.1–27.
31. Aida T, Chiyo K, Usami T, Ishikubo H, Imahashi R, Wada Y, Tanaka KF, Sakuma T, Yamamoto T, Tanaka K. Cloning-free CRISPR/Cas system facilitates functional cassette knock-in in mice. Genome Biol. 2015;16:87.
32. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F. Cpf1 Is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015;163:759–71.
33. Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, Minakhin L, Joung J, Konermann S, Severinov K, Zhang F, Koonin EV. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. Mol Cell. 2015;60:385–97.
34. Nakade S, Tsubota T, Sakane Y, Kume S, Sakamoto N, Obara M, Daimon T, Sezutsu H, Yamamoto T, Sakuma T, Suzuki KT. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. Nat Commun. 2014;5:5560.
35. Quadros RM, Harms DW, Ohtsuka M, Gurumurthy CB. Insertion of sequences at the original provirus integration site of mouse ROSA26 locus using the CRISPR/Cas9 system. FEBS Open Bio. 2015;5:191–7.
36. Miura H, Gurumurthy CB, Sato T, Sato M, Ohtsuka M. CRISPR/Cas9-based generation of knockdown mice by intronic insertion of artificial microRNA using longer single-stranded DNA. Sci Rep. 2015;5:12799.

Genome-Editing Technology in CRISPR/Cas System: How to Increase Knock-In Efficiency in Mouse Zygotes

Takuro Horii and Izuho Hatada

Abstract The use of clustered regularly interspaced short palindromic repeats (CRISPR) and RNA-guided Cas9 nucleases, known as the CRISPR/Cas system, represents a major technological advance in mammalian gene disruption. CRISPR/Cas enables genome editing by inducing targeted DNA double-strand breaks (DSBs) that are repaired by error-prone, nonhomologous end-joining (NHEJ), or homology-directed repair (HDR). This system has emerged as an effective tool for gene knockout via NHEJ; however, it remains inefficient for precise editing of genome sequences depending on HDR. Nevertheless, HDR-mediated gene editing is essential for conditional knockout, introduction of reporter genes, and precise point mutation in mice. Many studies have examined, for example, conditions of Cas9 and guide RNA (gRNA), methods of their introduction, and molecules to increase efficiency. In this review, we describe various methods for increasing the efficiency of editing in mouse zygotes.

Keywords CRISPR/Cas • Cas9 • sgRNA • Double-strand break • Nonhomologous end-joining • Homology-directed repair • Knock-in

Abbreviations

amiRNA	Artificial microRNA
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
DSB	Double-strand breaks
dsDNA	Double-stranded DNA
ESC	Embryonic stem cell
gRNA	Guide RNA

Author Contributions T. Horii and I. Hatada wrote the manuscript.

T. Horii • I. Hatada (✉)
Laboratory of Genome Science, Biosignal Genome Resource Center, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15 Showa-machi, Maebashi, Gunma 371-8512, Japan
e-mail: hatada@gunma-u.ac.jp

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_5

HDR	Homology-directed repair
indels	Insertions or deletions
iPSC	Induced pluripotent stem cells
NHEJ	Nonhomologous end-joining
ssODN	Single-stranded oligo-DNA
TALEN	Transcription activator-like effector nucleases
tracrRNA	Trans-activating crRNA
ZFN	Zinc-finger nucleases

Introduction

The mouse is the most widely used mammalian model system because its genome can be precisely modified as desired. Conventional gene-targeting methods via homologous recombination are generally used for mouse embryonic stem cells (ESCs) followed by germline chimera production. However, these technologies are not available in species other than mouse. The recent development of site-specific endonucleases for selective genome cleavage has been an important advancement in genome engineering. These technologies include zinc-finger nucleases (ZFN) [1], transcription activator-like effector nucleases (TALEN) [2], and the CRISPR/Cas system [3]. These technologies enabled rapid production of gene editing not only in mice but also in other species. Previously, it took more than 1 year to obtain a homozygous mutant mouse by the conventional method using ESCs; now, founder (F_0) mice can be obtained within a month in these technologies. In particular, the construction of CRISPR/Cas is far easier and more economical than other engineered endonucleases, making this the most popular genome-editing system.

CRISPR/Cas System

The CRISPR/Cas system was initially identified as an adaptive immune system in bacteria. This system requires Cas9 nuclease and two small RNAs, CRISPR RNA (crRNA), which guides the Cas9 complex to the target sequence, and trans-activating crRNA (tracrRNA), which binds to crRNA. Cas9 nuclease and tracrRNA–crRNA form a ribonucleoprotein complex and generate a double-stranded break (DSB) on target sites [4]. When it is applied as a genome-editing tool, the process can be simplified using an engineered guide RNA (gRNA) containing a hairpin that mimics the tracrRNA–crRNA complex [3]. When targeted DNA DSBs have been made by the Cas9–gRNA complex, they are repaired by the two major DNA damage-repair pathways, nonhomologous end-joining (NHEJ) [5, 6] or homology-directed repair (HDR) [7, 8]. Error-prone NHEJ-mediated repair, which occurs rapidly and preferentially, induces small insertions or deletions (indels) at the DSB site, resulting in

disruption of gene function by frameshift mutations. By contrast, in the presence of a single- or double-stranded DNA template containing regions of homology to the sequences flanking the DSB, mutant alleles with precise-point mutations or DNA inserts can be generated by HDR.

Currently, CRISPR-mediated gene knockout through the error-prone NHEJ pathway works efficiently. For example, the knockout efficiency of a protein-coding gene is 20–60 % in mouse ESCs and zygotes [9, 10]. By contrast, precise introduction of a point mutation or a sequence fragment directed by the HDR-mediated repair pathway has remained inefficient [9–11]. To increase editing efficiency, a screening process by cell sorting or drug selection is usually required to obtain correctly edited cells. However, this strategy is difficult to apply to zygotes in part because the number of available zygotes is very limited compared to that of cultured cells (at most 10^2 zygotes versus more than 10^7 cultured cells per experiment). Therefore, improving the efficiency of precise gene editing in zygotes has remained a major challenge. Here, we review studies that have sought to increase the efficiency of gene editing, especially via the HDR-mediated repair pathway.

Delivery System of Cas9 and gRNA

Microinjection into zygotes is the most widely used method for producing gene-edited mice with the CRISPR/Cas system. To generate a NHEJ-mediated mutants, two components must be introduced into zygotes: the Cas9 nuclease and a gRNA that is complementary to the target DNA. To generate knock-in mice via the HDR pathway, donor DNA must also be introduced into the target cell. Because many expression vectors can be easily produced from large-scale cultures of genetically modified bacteria, the simplest method for generating knockout mice is injection of the expression vectors for Cas9 and gRNA into the pronucleus because many expression vectors are easily produced from large-scale culture of gene-modified bacteria. In addition, many laboratories have established routine systems of DNA injection for production of transgenic mice. However, this injection method runs the risk of vector integration into the chromosomes. Mashiko and colleagues reported that 4.3 % (2/46) of mutant pups generated by the circular DNA injection method carried the hCas9 transgene [12, 13]. Although this was lower than the average transgenic efficiency of linearized DNA, 33.4 % (173/684, $N = 26$ constructs), undesired integration of the vectors into the chromosomes still occurred to some extent when the DNA injection method.

Therefore, injection of in vitro transcribed RNA is a better alternative, and this method is most widely used today. However, because the Cas9 RNA and gRNA function in different locations within the cell (the cytoplasm and pronucleus, respectively), it was necessary to optimize this method. Horii and colleagues compared three different injection techniques [14]: (1) injection of circular DNA into the pronucleus, (2) injection of RNA into the pronucleus, and (3) injection of RNA into the cytoplasm. The results revealed that injection of RNA into the cytoplasm is the

most efficient method in terms of the numbers of generated blastocysts and pups. In addition, injection of RNA into the cytoplasm yielded the best overall knockout efficiency.

For knock-in mice, it is necessary to consider where the donor DNA should be introduced. In general, transgenic mice are produced by pronuclear injection of DNA rather than cytoplasmic injection of RNA. To optimize the HDR-mediated knock-in method for the CRISPR/Cas system, Yang and colleagues compared three different injection methods [10]: (1) simultaneous injection of RNA and donor DNA into the cytoplasm (2) simultaneous injection of RNA and donor DNA into the pronucleus and (3) injection of RNA into the cytoplasm followed 2 h later by pronuclear injection of donor DNA. Simultaneous injection of all components into the cytoplasm yielded 9–19 % of targeted blastocysts. Thus, donor DNA can be transported to the nucleus even if it is initially injected into the cytoplasm. Similarly, the simultaneous injection of all components into the pronucleus yielded 9–18 % targeted blastocysts. By contrast, the two-step procedure yielded at most 3 % of targeted blastocysts. These results suggest that simultaneous injection of RNA and DNA into the cytoplasm or pronucleus is the most efficient procedure for achieving targeted insertion. In addition, some groups reported that simultaneous injection of RNA and DNA into both the pronucleus and cytoplasm of zygotes resulted in a relatively high yield of targeted embryos [15–18].

Gene-edited mice can also be generated by direct injection of Cas9 protein rather than the Cas9 mRNA. A recent study reported efficient generation of NHEJ-mediated knockout mice by direct delivery of the Cas9 protein–RNA complex [19]. The yield of mutant alleles obtained was dose-dependent, reaching 88 % at the highest dose by pronuclear injection and 71 % by cytoplasmic injection. Aida and colleagues reported that Cas9 protein is also useful for generating reporter knock-in mice [20]. First, they tried to generate knock-in mice by direct pronuclear delivery of a Cas9 protein–gRNA complex; however, they could not obtain any mice carrying a functional gene cassette. Next, they tried microinjection of Cas9 protein combined with chemically synthesized crRNA and tracrRNA instead of gRNA, which resulted in targeted insertion of the transgene with 45.5 % efficiency. Although it is not clear why crRNA–tracrRNA is superior to gRNA, the combination of crRNA–tracrRNA and Cas9 protein may be optimal.

Fujii and colleagues also compared the effectiveness of different types of gRNA (long and short) for generation of NHEJ-mediated knockout mice [21]. The short type of gRNA, which is more widely used, has Cas9-associated sequences of ~40 nt [22]; by contrast, long gRNA is ~80 nt in length. The efficiency of genome modification resulting from injection into zygotes is much higher for longer type.

In addition to the microinjection, electroporation enables efficient mRNA delivery into mouse and rat zygotes [23, 24]. Microinjection into zygotes requires special skill and is too time consuming for large-scale production of gene-edited mice. By contrast, electroporation is simple and easy for beginners. In addition, Hashimoto and Takemoto successfully generated HDR-mediated knock-ins using single-stranded oligo-DNA (ssODN) [23]; 36.4 % (4/11) of the resultant embryos carried the loxP on the target site.

Single-Stranded Oligo-DNA and Double-Stranded DNA as Donor Templates

Knock-in mice carrying single nucleotide substitutions combined with ssODN donors have been generated by several groups (Table 1). By contrast, far fewer studies have reported the successful production of knock-in mice carrying gene cassettes derived from double-stranded DNA (dsDNA) (Table 2). The ssODN repair template requires only short homology arms and is inserted at a very high efficiency, whereas the dsDNA-repair template requires long homology arms and is inserted with low efficiency [18, 25]. When small epitope tags (V5, HA, or FLAG) or loxP is introduced, ssODN-containing homology arms of 40–60 bp are sufficient [9, 10, 18, 26–28]. By contrast, in many cases, larger tags or fluorescent markers (GST, mCherry, or GFP) often require dsDNA templates with homology arms of 1–3 kb on either side of the DSB sites [10, 15, 17, 20].

dsDNA is used for large insertions due to limitations of imposed by the short length of ssODN length. The maximum length of commercially synthesized ssODN is ~200 nt. Miura and colleagues utilized a two-step method to synthesize longer ssODN molecules (~514 bp) and demonstrate that they efficiently serve as repair templates for CRISPR/Cas-mediated knock-in [16]. In this system, RNA was first synthesized from a DNA template, and reverse transcription of the RNA was then performed to generate ssODN. Using this approach, they demonstrated that artificial microRNA (amiRNA) cassettes against the exogenous eGFP or endogenous orthodenticle homeobox 2 (Otx2) genes could be efficiently targeted to a predetermined locus (44.4–83.3 %).

Molecules That Increase HDR Efficiency

According to previous reports, efficiencies of knockout via NHEJ and knock-in via HDR do not differ significantly (Tables 1 and 2). However, the true efficiency of HDR editing events must be lower than that of NHEJ editing because HDR is in competition with NHEJ. In our experiment, for example, the efficiency of NHEJ events was 100 %, but that of HDR events was around 30 % on the same locus (data not shown). Therefore, NHEJ inhibition or HDR promotion is necessary to increase HDR efficiency.

Recent work showed that a drug, SCR7 increased HDR efficiency at the expense of NHEJ [18, 27, 29, 30]. Specifically, NHEJ repair is further divided into two subclasses: (1) KU- and DNA ligase IV-dependent, or “canonical” (C-NHEJ); and (2) DNA ligase I- or ligase III-dependent alternative end-joining (a-EJ or alt-NHEJ) [31]. SCR7 is an inhibitor of DNA ligase IV, a key enzyme in the C-NHEJ pathway. Indeed, SCR7 increased the efficiency of HDR-mediated genome editing up to tenfold (5.8 % versus 56.2 %) in resultant pups [18]. Maruyama and colleagues also reported that SCR7 treatment increased the efficiency of HDR events up to 19 fold in mice for at least four genes (*Kell*, *Igkc*, *Os9*, *Sgms2*) [27].

Table 1 Efficiency of knock-in mouse using ssODN as donor DNA

Donor DNA	Target	Cas9	gRNA	Concentration (Cas9/gRNA/donor DNA ng/μL)	Introduction method	Sample stage		Knock-in efficiency	References
loxP	*Tet1*	mRNA	gRNA	100/50/100	Cyto	Blastocyst		6/9 (66.7 %)	[9]
loxP	*Tet2*	mRNA	gRNA	100/50/100	Cyto	Blastocyst		9/15 (60.0 %)	
2 loxP (Tet1 and Tet2)	*Tet1 and Tet2*	mRNA	gRNA	100/50 each/ 100 each	Cyto	Blastocyst	2 loxP	1/14 (7.1 %)	
V5 tag	*Sox2*	mRNA	gRNA	100/50/200	Cyto	Postnatal		12/35 (34.3 %)	[10]
2 loxP (Left and Right)	*Mecp2*	mRNA	gRNA x2	100/50 each/ 200 each	Cyto	Postnatal	2 loxP (L & R)	16/98 (16.3 %)	
							1 loxp (L)	45/98 (46.0 %)	
							1 loxP (R)	25/98 (25.5 %)	
Crygc correct gene	*Crygc*	mRNA	gRNA	50/20/20	Cyto	Postnatal	Oligo1	5/29 (17.2 %)	[28]
							Oligo2	4/27 (25.9 %)	
FLAG-tag	*Hprt*	D10A, mRNA	gRNA x2	100/10 each/?	Cyto	Postnatal		8/10 (80 %)	[26]
Dmd correct gene	*Dmd*	mRNA	gRNA	10/10/10	PN	Postnatal		1/29 (3.4 %)	
				10/10/10	PN + cyto			4/58 (6.9 %)	
				50/20/10	PN			0/14 (0 %)	
				50/20/10	PN + cyto			2/23 (8.9 %)	
Point mutation (CG to TA)	*Tex15*	mRNA	gRNA	50/50/100	PN + cyto		SCR7(−)	2/17 (5.8 %)	[18]
							SCR7(+)	9/16 (56.2 %)	
Silent mutation	*Cdk2*	mRNA	gRNA	50/50/100	PN + cyto			10/27 (44.4 %)	

LPETG	*Kell*	mRNA	gRNA	100/50/100	Cyto	Blastocyst	SCR7(−)	11/41 (26.8 %)	[27]
							SCR7(+)	35/59 (59.3 %)	
LPETG	*Igkc*	mRNA	gRNA	100/50/100	Cyto	Embryo (E10)	SCR7(−)	1/20 (5.0 %)	
							SCR7(+)	5/22 (22.7 %)	
Stop codon	*Os9*	mRNA	gRNA	100/50/100	Cyto	Embryo (E10)	SCR7(+)	5/12 (41.7 %)	
amiR-EGFP	*EGFP*	mRNA	gRNA	10/10/14–20	PN + cyto	Embryo (E13.5)		5/6 (83.3 %)	[16]
amiR-Otx	*eEF2*	mRNA	gRNA	10/10/14–20	PN + cyto	Embryo (E14.5)		4/6 (66.7 %)	
loxP-amiR-EGFP-loxP	*EGFP*	mRNA	gRNA	10/10/14–20	PN + cyto	Embryo (E14.5)		4/9 (44.4 %)	
loxP	*mCherry*	mRNA	gRNA	200/100/400	Electroporation	Embryo (E9)		4/11 (36.4 %)	[23]

PN pronuclear, *Cyto* cytoplasm

Table 2 Efficiency of knock-in mouse using double-stranded DNA (dsDNA) as donor DNA

Donor DNA	Target	Cas9	gRNA	Concentration (Cas9/gRNA/donorDNA ng/μL)	Introduction method	Sample stage		Knock-in efficiency	References
mCherry	*Nanog*	mRNA	gRNA	100/50/200	Cyto	Postnatal		7/86 (8.1%)	[10]
GFP	*Oct4*	mRNA	gRNA	100/50/200	Cyto	Postnatal		3/10 (30.0%)	
PGK EGFP pA	*Ramp2*	mRNA	gRNA	200/50/10	PN+cyto	Blastocyst		4/7 (57%)	[17]
Ispd flox	*Ispd*	D10A, mRNA	gRNA x2	10/5 each/10	PN+cyto	Postnatal		1/13 (7.7%)	[15]
TetO-FLEX-EGFP-polyA	*Actb*	Protein	crRNA: tracrRNA	30/0.61 pmol each/10	PN	Postnatal		5/11 (45.5%)	[20]
Stop casette	*Sgms2*	mRNA	gRNA	100/50/100	Cyto	Embryo (E10)	SCR7(+)	1/8 (12.5%)	[27]

PN pronuclear, *Cyto* cytoplasm

To increase the frequency of HDR events, Chu and colleagues used gene silencing to suppress the key NHEJ molecules KU70 and DNA ligase IV [29], resulting in a 4.5-fold increase in the efficiency of HDR. They also found that co-expression of the adenovirus 4 (Ad4) E1B55K and E4orf6 proteins, which mediate the ubiquitination and proteasomal degradation of DNA ligase IV, increased the efficiency of HDR up to eightfold in both human and mouse cell lines.

Yu and colleagues developed a reporter-based screening approach for high-throughput identification of chemical compounds that increase efficiency of precise genome editing via HDR [32]. Using this screening method, they identified two small molecules, L755507 and brefeldin A, that can increase HDR efficiency threefold for large fragment insertions and ninefold for point mutations. They have also found small molecules, azidothymidine (AZT) and trifluridine (TFT), that decrease the HDR efficiency. Except for SCR7, these molecules have not been examined in zygotes, but they could be used to increase HDR frequency. Other small molecules can effectively activate or block certain DNA-repair pathways [33, 34], but they have not been examined in the CRISPR/Cas system. These small molecules could be also used to modulate CRISPR-induced genome editing and DNA repair via the HDR pathway.

On the other hand, low-dose irradiation promotes gene targeting in human pluripotent stem cells [35]. Hatada and colleagues reported that limited low-dose irradiation using either γ-rays or X-rays exposure (0.4 Gy) significantly increased the HDR efficiency of the CRISPR/Cas system, possibly through induction of DNA-repair/recombination machinery such as the ataxia-telangiectasia mutated (ATM), histone H2A.X, and RAD51 proteins. We summarize the molecules that promote HDR events (Fig. 1).

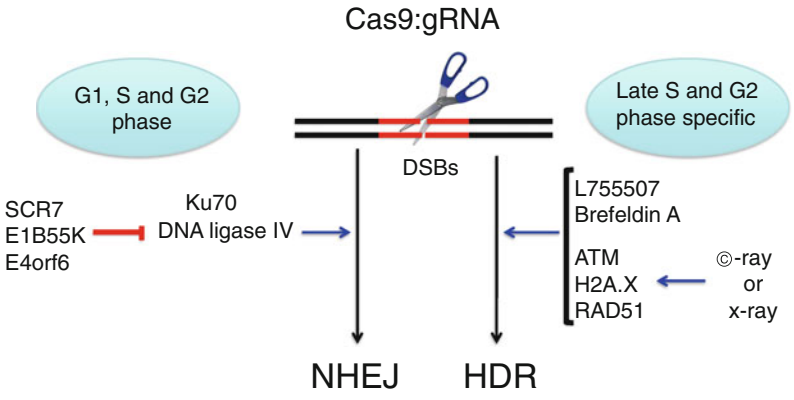

Fig. 1 Molecules that regulate homology-directed repair (HDR) and nonhomologous end-joining homology-directed repair NHEJ efficiency

Cell Cycle

NHEJ dominates DNA repair during the G_1-, S-, and G_2 phases, whereas HDR is restricted to late S- and G_2-phases (Fig. 1) when DNA replication is completed and sister chromatids are available to serve as repair templates [36]. However, the Cas9–gRNA complex is usually introduced into unsynchronized cells and zygotes. Lin and colleagues examined HDR efficiency depend on each cell cycle using synchronized cells [37]. The frequency of HDR events was increased dramatically relative to experiments in unsynchronized cells, with rates of HDR up to 38 % in HEK293T cells. In particular, treatment with nocodazole, which blocks cells at the M-phase, led to higher HDR efficiency.

On the other hand, synchronization of zygotes has not yet been reported. Zygotes at the pronuclear stage were classified into five pronuclear stages (PN1–PN5) according to pronuclear size and location in the cytoplasm [38]. The time at which Cas9 and gRNA are usually microinjected into pronuclei or cytoplasm during PN3 to PN5, which correspond to the S-phase to G_2-phase. Therefore, the cell cycle of zygotes could be almost adequate for HDR-mediated knock-in.

Perspective

HDR-mediated gene editing is has been applied to conditional knockouts, introduction of reporter genes, and precise point mutations in experimental animals. In addition, studies using mice and rats have shown that the CRISPR/Cas system can correct diseases caused by mutations and reverse their phenotypes [28, 39]. However, the final goal of gene correction would be therapeutic application in humans. Precise correction of mutant genes has already been reported in human-induced pluripotent stem cells (iPSC) [40–44], suggesting that HDR-mediated gene editing by CRISPR/Cas would be effective for gene therapy. Liang and colleagues reported gene correction by CRISPR/Cas using human tri-pronuclear zygotes [45]. Specifically, they found that the HDR efficiency of the endogenous human β-globin (HBB) gene was low, whereas CRISPR/Cas effectively cleaved the endogenous HBB gene. In addition, edited embryos exhibited mosaicism, and off-target cleavage also occurred. Thus, gene correction of human zygotes requires further improvement of the fidelity and specificity of the CRISPR/Cas system.

Conclusion

We described various methods and molecules that could enhance HDR frequency. Many molecules can increase the frequency of HDR events in cultured cells but most of them have not been examined in zygotes. In the future, a combination of

these molecules with optimized injection methods could increase the efficiency of knock-in by HDR-mediated repair.

Acknowledgments This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; the Ministry of Health, Labor, and Welfare of Japan; the National Institute of Biomedical Innovation; and the Takeda Science Foundation.

Conflicts of Interest The authors declare no conflicts of interest.

References

1. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A. 1996;93(3):1156–60.
2. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011;29(2):143–8. doi:10.1038/nbt.1755.
3. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816–21. doi:10.1126/science.1225829.
4. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471(7340):602–7. doi:10.1038/nature09886.
5. Barnes DE. Non-homologous end joining as a mechanism of DNA repair. Curr Biol. 2001;11(12):R455–7.
6. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010;79:181–211. doi:10.1146/annurev.biochem.052308.093131.
7. van den Bosch M, Lohman PH, Pastink A. DNA double-strand break repair by homologous recombination. Biol Chem. 2002;383(6):873–92. doi:10.1515/bc.2002.095.
8. Wyman C, Kanaar R. DNA double-strand break repair: all's well that ends well. Annu Rev Genet. 2006;40:363–83. doi:10.1146/annurev.genet.40.110405.090451.
9. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013;153(4):910–8. doi:10.1016/j.cell.2013.04.025.
10. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell. 2013;154(6):1370–9. doi:10.1016/j.cell.2013.08.022.
11. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823–6. doi:10.1126/science.1232033.
12. Mashiko D, Young SA, Muto M, Kato H, Nozawa K, Ogawa M, Noda T, Kim YJ, Satouh Y, Fujihara Y, Ikawa M. Feasibility for a large scale mouse mutagenesis by injecting CRISPR/Cas plasmid into zygotes. Dev Growth Differ. 2014;56(1):122–9. doi:10.1111/dgd.12113.
13. Mashiko D, Fujihara Y, Satouh Y, Miyata H, Isotani A, Ikawa M. Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. Sci Rep. 2013;3:3355. doi:10.1038/srep03355.
14. Horii T, Arai Y, Yamazaki M, Morita S, Kimura M, Itoh M, Abe Y, Hatada I. Validation of microinjection methods for generating knockout mice by CRISPR/Cas-mediated genome engineering. Sci Rep. 2014;4:4513. doi:10.1038/srep04513.

15. Lee AY, Lloyd KC. Conditional targeting of Ispd using paired Cas9 nickase and a single DNA template in mice. FEBS Open Bio. 2014;4:637–42. doi:10.1016/j.fob.2014.06.007. eCollection 2014.
16. Miura H, Gurumurthy CB, Sato T, Sato M, Ohtsuka M. CRISPR/Cas9-based generation of knockdown mice by intronic insertion of artificial microRNA using longer single-stranded DNA. Sci Rep. 2015;5:12799. doi:10.1038/srep12799.
17. Sakurai T, Watanabe S, Kamiyoshi A, Sato M, Shindo T. A single blastocyst assay optimized for detecting CRISPR/Cas9 system-induced indel mutations in mice. BMC Biotechnol. 2014;14:69. doi:10.1186/1472-6750-14-69.
18. Singh P, Schimenti JC, Bolcun-Filas E. A mouse geneticist's practical guide to CRISPR applications. Genetics. 2015;199(1):1–15. doi:10.1534/genetics.114.169771.
19. Sung YH, Kim JM, Kim HT, Lee J, Jeon J, Jin Y, Choi JH, Ban YH, Ha SJ, Kim CH, Lee HW, Kim JS. Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. Genome Res. 2014;24(1):125–31. doi:10.1101/gr.163394.113.
20. Aida T, Chiyo K, Usami T, Ishikubo H, Imahashi R, Wada Y, Tanaka KF, Sakuma T, Yamamoto T, Tanaka K. Cloning-free CRISPR/Cas system facilitates functional cassette knock-in in mice. Genome Biol. 2015;16:87. doi:10.1186/s13059-015-0653-x.
21. Fujii W, Kawasaki K, Sugiura K, Naito K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. Nucleic Acids Res. 2013;41(20), e187. doi:10.1093/nar/gkt772.
22. Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z, Zhang X, Zhang P, Huang X. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. Cell Res. 2013;23(5):720–3. doi:10.1038/cr.2013.46.
23. Hashimoto M, Takemoto T. Electroporation enables the efficient mRNA delivery into the mouse zygotes and facilitates CRISPR/Cas9-based genome editing. Sci Rep. 2015;5:11315. doi:10.1038/srep11315.
24. Kaneko T, Sakuma T, Yamamoto T, Mashimo T. Simple knockout by electroporation of engineered endonucleases into intact rat embryos. Sci Rep. 2014;4:6382. doi:10.1038/srep06382.
25. Yang H, Wang H, Jaenisch R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. Nat Protoc. 2014;9(8):1956–68. doi:10.1038/nprot.2014.134.
26. Fujii W, Onuma A, Sugiura K, Naito K. Efficient generation of genome-modified mice via offset-nicking by CRISPR/Cas system. Biochem Biophys Res Commun. 2014;445(4):791–4. doi:10.1016/j.bbrc.2014.01.141.
27. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol. 2015;33(5):538–42. doi:10.1038/nbt.3190.
28. Wu Y, Liang D, Wang Y, Bai M, Tang W, Bao S, Yan Z, Li D, Li J. Correction of a genetic disease in mouse via use of CRISPR-Cas9. Cell Stem Cell. 2013;13(6):659–62. doi:10.1016/j.stem.2013.10.016.
29. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, Kühn R. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol. 2015;33(5):543–8. doi:10.1038/nbt.3198.
30. Srivastava M, Nambiar M, Sharma S, Karki SS, Goldsmith G, Hegde M, Kumar S, Pandey M, Singh RK, Ray P, Natarajan R, Kelkar M, De A, Choudhary B, Raghavan SC. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. Cell. 2012;151(7):1474–87. doi:10.1016/j.cell.2012.11.054.
31. Bétermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? PLoS Genet. 2014;10(1), e1004086. doi:10.1371/journal.pgen.1004086.
32. Yu C, Liu Y, Ma T, Liu K, Xu S, Zhang Y, Liu H, La Russa M, Xie M, Ding S, Qi LS. Small molecules enhance CRISPR genome editing in pluripotent stem cells. Cell Stem Cell. 2015;16(2):142–7. doi:10.1016/j.stem.2015.01.003.
33. Hollick JJ, Golding BT, Hardcastle IR, Martin N, Richardson C, Rigoreau LJ, Smith GC, Griffin RJ. 2,6-Disubstituted pyran-4-one and thiopyran-4-one inhibitors of DNA-dependent protein kinase (DNA-PK). Bioorg Med Chem Lett. 2003;13(18):3083–6.

34. Rahman SH, Bobis-Wozowicz S, Chatterjee D, Gellhaus K, Pars K, Heilbronn R, Jacobs R, Cathomen T. The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. Hum Gene Ther. 2013;24(1):67–77. doi:10.1089/hum.2012.168. Epub 2012 Dec 10.
35. Hatada S, Subramanian A, Mandefro B, Ren S, Kim HW, Tang J, Funari V, Baloh RH, Sareen D, Arumugaswami V, Svendsen CN. Low-dose irradiation enhances gene targeting in human pluripotent stem cells. Stem Cells Transl Med. 2015;4(9):998–1010. doi:10.5966/sctm.2015-0050.
36. Heyer WD, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. Annu Rev Genet. 2010;44:113–39. doi:10.1146/annurev-genet-051710-150955.
37. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. elife. 2014;3:e04766. doi:10.7554/eLife.04766.
38. Adenot PG, Mercier Y, Renard JP, Thompson EM. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. Development. 1997;124(22):4615–25.
39. Yoshimi K, Kaneko T, Voigt B, Mashimo T. Allele-specific genome editing and correction of disease-associated phenotypes in rats using the CRISPR-Cas platform. Nat Commun. 2014;5:4240. doi:10.1038/ncomms5240.
40. Huang X, Wang Y, Yan W, Smith C, Ye Z, Wang J, Gao Y, Mendelsohn L, Cheng L. Production of gene-corrected adult beta globin protein in human erythrocytes differentiated from patient iPSCs after genome editing of the sickle point mutation. Stem Cells. 2015;33(5):1470–9. doi:10.1002/stem.1969.
41. Li HL, Fujimoto N, Sasakawa N, Shirai S, Ohkame T, Sakuma T, Tanaka M, Amano N, Watanabe A, Sakurai H, Yamamoto T, Yamanaka S, Hotta A. Precise correction of the dystrophin gene in Duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. Stem Cell Rep. 2015;4(1):143–54. doi:10.1016/j.stemcr.2014.10.013.
42. Ousterout DG, Kabadi AM, Thakore PI, Majoros WH, Reddy TE, Gersbach CA. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. Nat Commun. 2015;6:6244. doi:10.1038/ncomms7244.
43. Song B, Fan Y, He W, Zhu D, Niu X, Wang D, Ou Z, Luo M, Sun X. Improved hematopoietic differentiation efficiency of gene-corrected beta-thalassemia induced pluripotent stem cells by CRISPR/Cas9 system. Stem Cells Dev. 2015;24(9):1053–65. doi:10.1089/scd.2014.0347.
44. Xie F, Ye L, Chang JC, Beyer AI, Wang J, Muench MO, Kan YW. Seamless gene correction of β-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. Genome Res. 2014;24(9):1526–33. doi:10.1101/gr.173427.114.
45. Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z, Lv J, Xie X, Chen Y, Li Y, Sun Y, Bai Y, Songyang Z, Ma W, Zhou C, Huang J. CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. Protein Cell. 2015;6(5):363–72. doi:10.1007/s13238-015-0153-5.

Developments in the Generation of Reporter Stem Cells

Samvel Avagyan, Kyle Sylakowski, and Uma Lakshmipathy

Abstract Pluripotent stem cells and their ability to form any cell type in the human body has allowed researchers to use them as research tools for applications in drug screening, basic developmental research, and potential therapeutic implementations. An emerging need in stem cell biology is efficient and homogeneous differentiation of stem cells into mature, specialized, functional cells.

Reporter stem cell lines are valuable models that enable noninvasive, live monitoring of marker onset and expression in a cell-specific manner. Several methods have been used to derive such cell lines based on lineage promoter-driven reporter expression. A more regulated expression achieved with a reporter knock-in into the endogenous promoter loci was less utilized because of the associated technical difficulty; however, new advances in genome-editing technologies has lowered these barriers for creating knock-in reporter lines. This chapter provides an overview of the methodology and potential applications of reporter stem cell lines.

Keywords Reporter lines • Pluripotent stem cells • Cell models • Gene knock-in • Cell engineering

S. Avagyan
Life Sciences Solutions, Thermo Fisher Scientific,
5781 Van Allen Way, Carlsbad, CA 92008, USA

CIRM Bridges Program, San Diego State University, San Diego, CA, USA

Rensselaer Polytechnic Institute, Troy, NY, USA

K. Sylakowski
Life Sciences Solutions, Thermo Fisher Scientific,
5781 Van Allen Way, Carlsbad, CA 92008, USA

University of Pittsburgh, Pittsburgh, PA, USA

U. Lakshmipathy (✉)
Life Sciences Solutions, Thermo Fisher Scientific,
5781 Van Allen Way, Carlsbad, CA 92008, USA
e-mail: uma.lakshmipathy@Thermofisher.com

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_6

Introduction

Human pluripotent stem cells (PSCs) such as embryonic stem cells (hESC) and induced pluripotent stem cells (iPSC) hold great promise in today's world of regenerative medicine. Their ability to self-renew and differentiate into any specialized cell in the body [1, 2] has given researchers in the biomedical field a unique tool to closer translate their research into the clinic. PSCs have the potential to be used in a variety of different research facets: from re-growing limbs, to conserving endangered animals, or as a cell therapy for replacing dead or dying cells. However, one of the most beneficial and most encouraging uses for PSCs is their ability to create in vitro human cell models.

A key challenge in realizing the full potential of pluripotent stem cells is the lack of robust methods to track and monitor cell behavior. The availability of stable modified stem cells, expressing linage-specific reporters, would present an ideal platform for the real-time tracking and monitoring of cells. Such reporter lines facilitate visualization and or enrichment based on the reporter gene used, as cells transition from a pluripotent state to specific differentiated states [3]. Lineage tracking enables development of optimal methods for differentiation and provides a live cell method to monitor cell fate behaviors in normal and diseased state.

Traditionally, reporter constructs have been composed of minimal, lineage-specific, promoter-driven reporters such as fluorescent genes or drug-selectable markers, and have been delivered using random or site-specific integrational methods. To overcome the constraints of minimal promoter specificity, large constructs carrying the full regulatory elements have been devised. However, the most context-specific method is insertion of reporters at the endogenous promoter loci to generate promoter knock-in lines. Genome-editing methods mediated via homologous recombination (HR) has been used to create disease models through gene targeting and modification [1–4]. These methods also enable targeted insertion of the reporter into promoter regions that are expressed in a lineage-specific manner such that expression is turned on only in the differentiated cell type [5]. This chapter highlights the methods used for generation of reporter pluripotent stem cell lines.

Gene Reporter Pluripotent Lines as Cell Models

Antibody staining or reverse transcriptase-polymer chain reaction (RT-PCR) methods specific for antigen or gene markers specific for that lineage traditionally achieve detection of specific cell type in a heterogeneous mixture [4–6]. These methods require a significant number of cells and are typically end-point assays that do not allow real-time monitoring of cells. Additionally, enrichment of terminally differentiated cells or elimination of undifferentiated cells is critical for therapeutic application where residual pluripotent stem cells can lead to teratoma formation in vivo [7].

Table 1 Commonly used reporters for visualization and enrichment of lineage-specific cell types

Promoter	Reporter		
Method	Gene mMarker	Advantages	Disadvantages
Exogenous lineage promoter or endogenous promoter Loci	Gene tags: bLac, bGal, AP	Enzymatic methods and hence signal can be amplified	Cells need to be fixed for visualization
	Fluorescent proteins: GFP/CFP/YFP, DsRED	Enables visualization of cells	High-level expression is required to be visualized over autofluorescence
	Bioluminescent marker: luciferase	Detection of sensitivity and use in biochemical assays	Requires cofactors/substrates and specialized equipment
	Drug selectable markers: Hyg, Neo, Puro, BSD	Sensitive even at low expression levels	Poor survival at low density, does not allow visualization
	Suicide genes: HSV-TK, diphtheria toxin-A	Allows enrichment or elimination of desired cells	Poor survival at low density, does not allow visualization

Gene reporters are a valuable in vitro tool that facilitates live monitoring and tracking a cell type of interest [3]. Lineage reporters are either created using minimal lineage-specific promoter-driven reporter systems or via knock-in of the reporter into the endogenous promoter region. Choice of reporter is critical based on intended application. Although there is a wide choice of reporters that can be used, each system offers its own advantages and disadvantages (Table 1).

Green fluorescent protein (GFP) and variants, such as cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and DsRed, have been popularly used to visualize cells. Their in vivo safety and applicability has been well demonstrated in mouse systems [8, 9]. However, even with enhanced versions of GFP, 10^5 copies per typical mammalian cells is required for visualization of fluorescence. Therefore, despite the availability of sensitive instruments for signal detection, the lower sensitivity threshold is actually set by expression level and cellular autofluorescence [10]. In contrast, gene tags such as β-lactamase, when used with membrane-permeant ester substrates, can be detected at low levels (~50 molecules per cells.) [11]. Bioluminescence reporters offer a distinct advantage for biochemical assays because of their detection sensitivity and emission duration times [12]. Drug resistance markers offer an alternate option wherein intermediate copy numbers of the reporter are sufficient for enrichment of the desired cell population and have been used extensively [13, 14]. The major disadvantage of this approach is the need for optimization of drug for each target cell type. Also, it does not allow visual monitoring of marker expression. However, a combination of one or more of these reporter systems can be used to successfully build robust tracking and enrichment reporter cell lines. A combination of fluorescent

reporter with drug resistance gene is most commonly used. However, the difference in their relative sensitivities, especially with weak promoters, can result in drug-resistant cells that do not show visible fluorescence [14].

Methods for the Generation of Reporter Pluripotent Stem Cells

There are two main ways to create reporter lines: the first is by inducing lineage-specific promoter-directed reporter expression; and the second is by insertion of the reporter gene into the endogenous promoter locus.

Minimal Promoter-Driven Reporter Expression

Exogenous lineage-specific promoter and regulatory fragments can be cloned upstream of a reporter gene to generate cell-specific promoter reporters. In most cases, DNA sequences upstream of the gene-encoding regions containing cis-acting elements that are conserved across species are designated as the promoter region; this represents a very restrictive representation of the regulatory elements as in some cases gene regulation is known to be complex, requiring combination of proximal and distal elements [15]. Consequently, the primary concern with the use of promoter-reporter systems is that the reporter may fail to faithfully recapitulate the activity of the endogenous target gene. It is therefore essential to carry out extensive validation of the promoter-reporter constructs to confirm context-specific expression of the reporter.

Expression level is also largely dependent on how the construct is delivered for stable expression into host cells. The promoter-reporter construct can either be maintained episomally without integration, or integrated either randomly or site specifically into the host genome (Fig. 1).

Nonintegrational Methods

Epstein–Barr virus (EBV)-based episomal vectors have been successfully used to stably express the gene of interest in multiple types of cells without integrating into the host genome [16]. This system offers an appealing alternative because it is relatively free from chromosomal effects associated with genomic integration methods and has been used to create pluripotent reporter lines with Pou5F1 (Oct4) promoter [17]. EBV vectors are rather large in size (more than 10 kb) and, in the absence of an efficient transfection method, this method can be limiting.

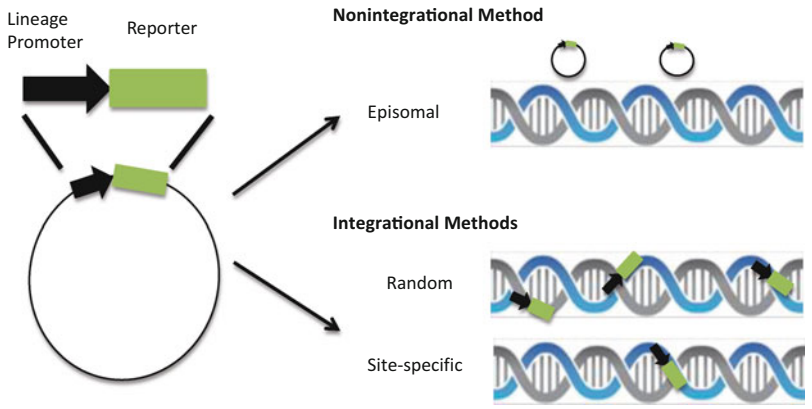

Fig. 1 Methods for generation of lineage reporter lines using lineage promoter-driven reporter constructs. Nonintegrational methods using episomal vectors prevent genomic integration of the transgene, whereas integration methods result in either random insertion or site-specific integration of the transgene into the host genome

Integrational Methods

Several viral and DNA-based methods facilitate the integration of delivered DNA constructs into the host genome, either in a random manner or to specific loci in the genome.

Random integration of the promoter-reporter construct can be achieved via naked DNA delivery or as lentiviral particles. In both cases, high copy numbers of the construct can be achieved with efficient delivery into cells. However, because the expression is largely dependent on the site of integration and hence subject to chromatin dynamics and epigenetic regulation, gene reporter expression is often diminished or silenced [18]. In addition, ESCs transduced with a high viral load of GFP transgene have been shown to result in reduced differentiation capability [19].

Site-specific integration of the promoter-reporter construct into a safe harbor site minimizes genomic loci effects to a large degree. Recombination-mediated cassette exchange has been shown to be effective to genetically modify the transcriptionally active ROSA26 locus in human ESC [20]. PhiC31 integrase mediates targeting of transgenes to specific hotspots and has been successfully used to generate constitutive and lineage-specific expression targeted to chromosome 13 in human ESC [21, 22]. Other safe-harbor genomic loci that are shown to support sustained transgene expression in embryonic stem cells are ENVY [23] and AAVS1 [24]. A major limitation with site-specific insertion of lineage-specific promoter-driven reporter is that one or two copies of the gene may not be sufficient to facilitate robust expression of the reporter. The best way to overcome constraints with promoter-reporter construct systems is the development of promoter knock-in reporter lines.

Reporter Knock-In into Endogenous Promoter Site

Reporter genes can be inserted into specific genomic sites of interest via homologous recombination [25]. The targeting construct composed of a core region carrying the reporter cassette is flanked by homology arms that recombine with the target genomic loci, resulting in site-specific insertion of the reporter [26]. This process is fairly inefficient, generally occurring at a rate of one in a million cells that can be significantly enhanced in the presence of double-strand breaks [27]. New gene-targeting technologies aim to precisely cleave genomic loci to facilitate insertion or deletion of genes at the specific cleavage site (Fig. 2). Several reporter lines in ESC and iPSC have been generated using these methods (Table 2).

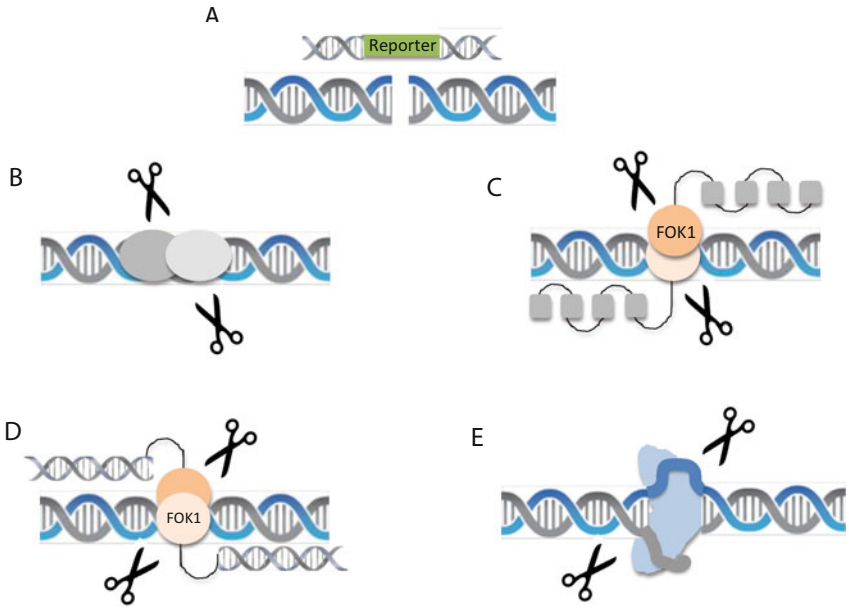

Fig. 2 Gene-editing methods for generation of reporter knock-in into endogenous promoter sites. **a** Traditional methods relied on homologous recombination between the target genomic site at the endogenous promoter locus and homology arm flanking the reporter. Meganucleases that use engineered versions of naturally occurring restriction enzymes with extended recognition sequences comprise both the DNA recognition and cleavage on a single domain (**b**). The newer systems ZFN (**c**), TALENS (**d**), and CRISPR (**e**) rely on engineered molecular scissors that precisely create double-strand breaks at specific genomic sites. ZFN and TALENS both rely on proteins containing a DNA-cleaving enzyme and a DNA-binding region that specifically recognizes the genomic region of interest. CRISPR relies on a DNA-cleaving protein guided by an RNA molecule that is complementary to the target genomic region of interest

Table 2 Examples of reporter knock-in pluripotent lines generated using different gene-editing methods

Cell type	Locus	Reporter	Method	Cell model	References
Human ESC, human iPSC	POU5F1 (Oct4)	GFP	HR, ZFN	Pluripotency	[24, 28]
HUES1, HUES3, KhES3 human ESC, 246H1 & 243H7 human iPSC	NANOG	GFP, Neomycin	HR, AAV	Pluripotency	[29, 30]
Human ESC	NESTIN	GFP	HR	Neuroepithelium	[31]
HUES 9 human ESC	FEZf2	YFP	HR	Corticospinal neuron	[32]
HES3, MEL1. H9 human ESC, DF19-9-7T human iPSC	NKX2.1	GFP	HR	Forebrain, neuronal	[33]
BG01 human ESC	Olig-2	GFP	HR	Oligodendrocyte	[34]
KhES3 human ESC, 246H1 human iPSC	HB9	GFP	HDAdVs	Motor neuron	[35]
Human iPSC	GFAP	GFP	ZFN	Astrocyte	[36]
Human ESC, human iPSC	PITX3	GFP	ZFN	Dopamine neuron	[24]
Human ESC	MIXL1	GFP	HR	Primitive streak, hematopoietic precursor	[37]
Human ESC	NKX2.5	GFP	HR	Cardiac progenitor	[38]
Human ESC	MESP1, NKX2.5	mCherry/eGFP	HR	Cardiac progenitor	[39]

Homologous Recombination

Since the idea of double-strand break (DSB) repair by HR was described [40], this method has provided a powerful method for creating cell models in human stem cells. More recently, this method has been employed to generate engineered human ESCs. In 2003 Zwaka and Thomson performed targeted gene editing in two loci of human ESC, HPRT-1 and POUF1 [28]. Shortly thereafter, other investigators have successfully targeted various loci including the oligodendrocyte lineage-specific Olig2 gene and the safe-harbor hROSA26 loci [20, 34].

Despite these studies, the traditional method of HR-mediated genome engineering has shortfalls. Given the low efficiency of success, the method requires millions of cells [20, 28, 34, 41]. Furthermore, successful transfection and clonal isolation of modified ESCs is a challenge because these cells prefer to remain as clumps [28]. Thus, the next milestone in genome editing would be developing more efficient methods of HR induction. These methods rely on recognition of specific DNA sequences to induce double-strand breaks adjacent to the chromosomal target site, thus facilitating higher recombinogenic events.

Mega Nucleases

Specific DNA-binding proteins are selected from a mutant library of restriction enzymes to generate site-specific cleavage of DNA for improved homologous recombination efficiency and reduced length of genomic DNA homology [42]. This approach requires a significant timeline for development of the tools (~8 months) and to date, versatility has not been demonstrated.

Zinc-Finger Nucleases

Zinc-finger nucleases include a nonspecific Folk1 cleavage domain fused to zinc-finger proteins that are composed of three to four zinc-finger motifs with each motif specifically recognizing a nucleotide triplet. The modular assembly approach involves the use of a preselected library of zinc-finger modules selected either via rational design or from a large combinatorial library [43, 44]. Several cell types has been successfully targeted using ZFN-mediated gene targeting [45–47]. The first study to use a lenti-based delivery of a ZFN expression cassette and the donor construct into human ESC reported a high gene-editing rate, even in the absence of selection [45]. Subsequent studies successfully used plasmid DNA in human ESC and iPSC for targeting drug resistance reporters to specific genes [41], and in particular for the creation of lineage-specific reporter lines via targeting of endogenous loci such as Oct4 (Pou5F1), PITX3, and GFAP [24, 36].

TALEN

Transcription activator-like effector nucleases, or TALENs, are a genome-modifying technology that allows highly specific double-stranded breaks at any DNA sequence. The specificity is achieved by recognition of a central repetitive region of 33–35 amino acids with two variable amino acids called the repeat variable diresidues (RVD). TALENs are designed in pairs using a combination of specific DNA-binding proteins called transcription activator-like effector proteins (TALEs) and an artificial nuclease or restriction enzyme such as Fok1 nuclease to bind to opposing DNA target sites for the generation of double-strand breaks [48]. This method has been successfully used for the knock-in of endogenous Oct4 and PITX3 promoters in ESC and iPSC with frequencies similar to that observed with ZFN technology [24, 49]. The combination of high targeting efficiency and easier design, facilitated by the simpler recognition code, makes this a very powerful tool for stable modification of pluripotent stem cells.

CRISPR

The most recent advent in genome editing has been the development of clustered regularly interspaced palindromic repeat (CRISPR)-guided Cas9 nuclease-mediated cleavage. There is a major aspect of CRISPR/Cas9 that makes it unique. The Cas9 nuclease remains independent of the guide sequence until the two are introduced upon performing an experiment. Having this flexibility means that targeting multiple sites is relatively simple; as the CIRSPR/Cas9 method only requires the synthesis of only a new guide RNA, instead of the production of a new target-specific nuclease from scratch [50]. The guide RNA is a single transcript consisting of two wild-type parts: target-specific crRNA and tracrRNA that helps to properly dock the Cas9 nuclease [51]. Cleavage occurs following binding of this RNA to its complementary strand of a 17- to 20-nucleotide-long target sequence upstream of the 3-nucleotide NGG protospacer adjacent domain (PAM). The biggest advantage of CRISPR-Cas9 is the specificity, efficiency, and ease with which the system can be utilized to target multiple sites, individually or together [52, 53].

Alternate Approaches

The methods described in earlier sections rely on stable modification of cells, which adds time and considerable effort to the generation and validation of stable clones. Recently, Miki et al. reported the use of synthetic microRNA (miRNA) switches for the detection and purification of specific cell populations [54]. The

method utilizes identification and purification based on endogenous miRNA activity. A cell type-specific miRNA is tagged with a reporter such as a fluorescent protein and delivered as a synthetic mRNA. Using the appropriate miRNA switch, cardiomyocytes, hepatocytes, and insulin-producing cells were successfully purified from differentiating human iPSC cultures [54]. This method provides a powerful approach that does not rely on DNA but is a safe and rapidly degradable method. The disadvantage to this is the fact that the microRNAs are short lived and the switches cannot be used in applications that require monitoring over long periods of time.

Potential Uses of Reporter Stem Cells

The applicability of reporter lines in pluripotent workflow is summarized in Fig. 3. Reporters can either be expressed in the pluripotent state and turned off with differentiation or vice versa, where the reporter expression is off to begin with and is turned on with differentiation. Reporter lines enable isolation of a homogeneous population of cells that is critical for characterization and for use in downstream applications. In addition, a visual reporter offers the opportunity of tracking cell fate progression in live cells.

In the case of pluripotent reporters, creation and characterization of the reporter line is relatively easy because clones can be derived based on expression of the reporter [17, 22, 24, 28–30]. A pluripotent reporter line can be used as a model to identify culture and maintenance conditions that preserve the pluripotent state. In addition, presence of residual undifferentiated cells following differentiation can be monitored and eliminated. Although engineered stem cells pose a higher barrier for transplantation, these cell models provide a valuable tool to identify optimal workflow solutions.

Lineage reporters, on the other hand, rely on onset of reporter expression as the pluripotent cell undergoes differentiation into lineage of choice; this has been best showcased in cardiomyocyte differentiation to help and define cardiac cell lineages [55, 56]. Fluorescent reporters and drug-selectable markers driven by cardiac and/or subtype-specific promoters such as alpha-MHC or NCX-1 promoter have been used for the enrichment of cardiac cell types to purity of greater than 90 % [14, 57–59]. However, as both alpha MHC and NCX-1 are expressed in all cardiomyocytes, the resulting cells were a mixture of atrial, ventricular, and nodal cells. This approach was further utilized for the specific enrichment of the ventricular cardiac subtype in ventricular myosin light chain 2v (MLC2V)-eGFP reporter [60] and nodal-like cells using cGATA6-GFP reporter [61]. More recently, knock-in human ESC lines MESP1-mCherry, NKX2.5-GFP, and a dual cardiac reporter carrying both these reporters has facilitated monitoring differentiation of hESC to precardiac mesoderm to cardiac lineage [38, 56]. Interestingly, a novel approach of miRNA switches was recently used to purify ESC-derived cardiomyocytes. Here, a synthetic mRNA

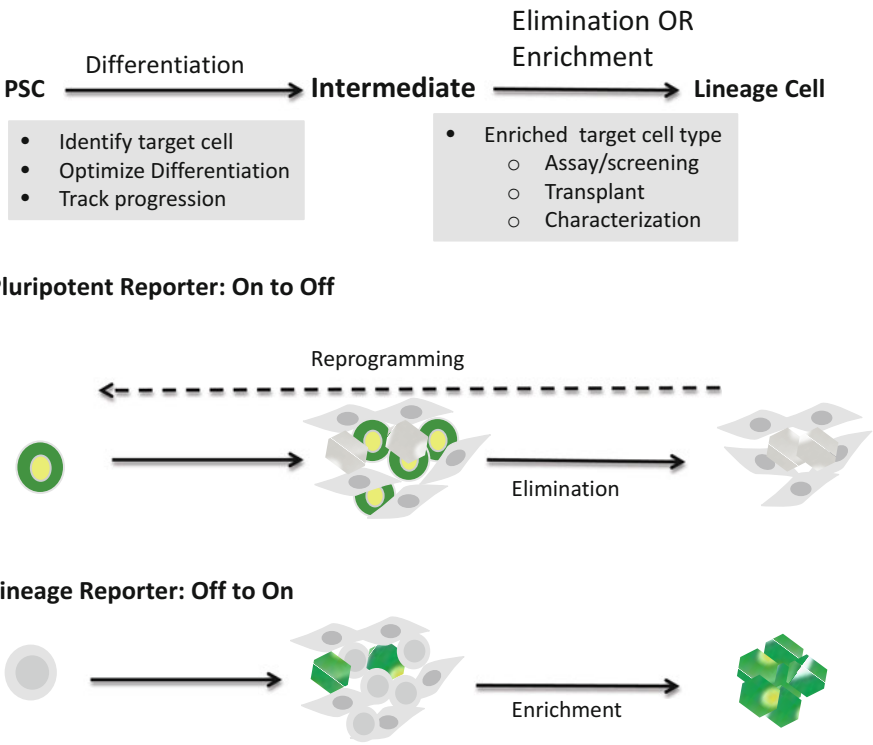

Fig. 3 Pluripotent and lineage reporters and their utility in downstream applications. Cell type of interest (*green*) can be either enriched or eliminated isolated from other cells (*gray*) in a heterogeneous mixture of cells based on reporter expression

composed of cardiomyocyte-specific miRNA driving expression of a selection reporter achieved cardiomyocytes with greater than 90 % purity; the resulting cells further survived and were successfully engrafted in immunocompromised mice [54]. Nevertheless, these methods have provided great value for optimization of differentiation methods. A similar approach has been used for tracking and identifying the various intermediate stages [31–33] and neural subcell types [34–36] derived from pluripotent stem cells.

In summary, lineage reporters are valuable cell models for tracing and enriching target cell types of interest. With progress in gene-editing methods, one or more reporters can be engineered in the same pluripotent stem cell line and could provide visual cues for the progression from stem cell to progenitor to highly functional mature cell types. This approach is no longer limited by gene-editing methods but rather uses differentiation protocols that can derive mature cell types which express the marker of interest at a high enough levels. These editing technologies for the creation of multicolor reporter lines can serve as key cell models for optimization and development of methodologies to generate highly purified, mature, and functional cell types of interest.

References

1. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. Science. 1998;282(5391):1145–7.
2. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007;131(5):861–72.
3. Giudice A, Trounson A. Genetic modification of human embryonic stem cells for derivation of target cells. Cell Stem Cell. 2008;2(5):422–33.
4. Cai J, Zhao Y, Liu Y, Ye F, Song Z, Qin H, Meng S, Chen Y, Zhou R, Song X, et al. Directed differentiation of human embryonic stem cells into functional hepatic cells. Hepatology. 2007;45(5):1229–39.
5. Shim JH, Kim SE, Woo DH, Kim SK, Oh CH, McKay R, Kim JH. Directed differentiation of human embryonic stem cells towards a pancreatic cell fate. Diabetologia. 2007;50(6):1228–38.
6. Wang J, Zhao HP, Lin G, Xie CQ, Nie DS, Wang QR, Lu GX. In vitro hematopoietic differentiation of human embryonic stem cells induced by co-culture with human bone marrow stromal cells and low dose cytokines. Cell Biol Int. 2005;29(8):654–61.
7. Hentze H, Graichen R, Colman A. Cell therapy and the safety of embryonic stem cell-derived grafts. Trends Biotechnol. 2007;25(1):24–32.
8. Long JZ, Lackan CS, Hadjantonakis AK. Genetic and spectrally distinct in vivo imaging: embryonic stem cells and mice with widespread expression of a monomeric red fluorescent protein. BMC Biotechnol. 2005;5:20.
9. Hadjantonakis AK, Macmaster S, Nagy A. Embryonic stem cells and mice expressing different GFP variants for multiple non-invasive reporter usage within a single animal. BMC Biotechnol. 2002;2:11.
10. Tsien RY. The green fluorescent protein. Annu Rev Biochem. 1998;67:509–44.
11. Zlokarnik G, Negulescu PA, Knapp TE, Mere L, Burres N, Feng L, Whitney M, Roemer K, Tsien RY. Quantitation of transcription and clonal selection of single living cells with beta-lactamase as reporter. Science. 1998;279(5347):84–8.
12. de Almeida PE, van Rappard JR, Wu JC. In vivo bioluminescence for tracking cell fate and function. Am J Physiol Heart Circ Physiol. 2011;301(3):H663–71.
13. Wang D, Haviland DL, Burns AR, Zsigmond E, Wetsel RA. A pure population of lung alveolar epithelial type II cells derived from human embryonic stem cells. Proc Natl Acad Sci U S A. 2007;104(11):4449–54.
14. Anderson D, Self T, Mellor IR, Goh G, Hill SJ, Denning C. Transgenic enrichment of cardiomyocytes from human embryonic stem cells. Mol Ther. 2007;15(11):2027–36.
15. van Weerd JH, Badi I, van den Boogaard M, Stefanovic S, van de Werken HJ, Gomez-Velazquez M, Badia-Careaga C, Manzanares M, de Laat W, Barnett P, et al. A large permissive regulatory domain exclusively controls Tbx3 expression in the cardiac conduction system. Circ Res. 2014;115(4):432–41.
16. Yates JL, Warren N, Sugden B. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature. 1985;313(6005):812–5.
17. Thyagarajan B, Scheyhing K, Xue H, Fontes A, Chesnut J, Rao M, Lakshmipathy U. A single EBV-based vector for stable episomal maintenance and expression of GFP in human embryonic stem cells. Regen Med. 2009;4(2):239–50.
18. Xia X, Zhang Y, Zieth CR, Zhang SC. Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. Stem Cells Dev. 2007;16(1):167–76.
19. Clements MO, Godfrey A, Crossley J, Wilson SJ, Takeuchi Y, Boshoff C. Lentiviral manipulation of gene expression in human adult and embryonic stem cells. Tissue Eng. 2006;12(7):1741–51.
20. Irion S, Luche H, Gadue P, Fehling HJ, Kennedy M, Keller G. Identification and targeting of the ROSA26 locus in human embryonic stem cells. Nat Biotechnol. 2007;25(12):1477–82.

21. Thyagarajan B, Liu Y, Shin S, Lakshmipathy U, Scheyhing K, Xue H, Ellerstrom C, Strehl R, Hyllner J, Rao MS, et al. Creation of engineered human embryonic stem cell lines using phiC31 integrase. Stem Cells. 2008;26(1):119–26.
22. Macarthur CC, Xue H, Van Hoof D, Lieu PT, Dudas M, Fontes A, Swistowski A, Touboul T, Seerke R, Laurent LC, et al. Chromatin insulator elements block transgene silencing in engineered human embryonic stem cell lines at a defined chromosome 13 locus. Stem Cells Dev. 2012;21(2):191–205.
23. Costa M, Dottori M, Ng E, Hawes SM, Sourris K, Jamshidi P, Pera MF, Elefanty AG, Stanley EG. The hESC line Envy expresses high levels of GFP in all differentiated progeny. Nat Methods. 2005;2(4):259–60.
24. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, Katibah GE, Amora R, Boydston EA, Zeitler B, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol. 2009;27(9):851–7.
25. Davis RP, Grandela C, Sourris K, Hatzistavrou T, Dottori M, Elefanty AG, Stanley EG, Costa M. Generation of human embryonic stem cell reporter knock-in lines by homologous recombination. Curr Protoc Stem Cell Biol. 2009;Chap. 5:Unit 5B 1 1 1-34.
26. Fontes A, Lakshmipathy U. Advances in genetic modification of pluripotent stem cells. Biotechnol Adv. 2013;31(7):994–1001.
27. Jasin M. Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet. 1996;12(6):224–8.
28. Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. Nat Biotechnol. 2003;21(3):319–21.
29. Mitsui K, Suzuki K, Aizawa E, Kawase E, Suemori H, Nakatsuji N, Mitani K. Gene targeting in human pluripotent stem cells with adeno-associated virus vectors. Biochem Biophys Res Commun. 2009;388(4):711–7.
30. Fischer Y, Ganic E, Ameri J, Xian X, Johannesson M, Semb H. NANOG reporter cell lines generated by gene targeting in human embryonic stem cells. PLoS One 2010, 5(9). pii: e12533.
31. Noisa P, Urrutikoetxea-Uriguen A, Li M, Cui W. Generation of human embryonic stem cell reporter lines expressing GFP specifically in neural progenitors. Stem Cell Rev. 2010;6(3):438–49.
32. Ruby KM, Zheng B. Gene targeting in a HUES line of human embryonic stem cells via electroporation. Stem Cells. 2009;27(7):1496–506.
33. Goulburn AL, Alden D, Davis RP, Micallef SJ, Ng ES, Yu QC, Lim SM, Soh CL, Elliott DA, Hatzistavrou T, et al. A targeted NKX2.1 human embryonic stem cell reporter line enables identification of human basal forebrain derivatives. Stem Cells. 2011;29(3):462–73.
34. Xue H, Wu S, Papadeas ST, Spusta S, Swistowska AM, MacArthur CC, Mattson MP, Maragakis NJ, Capecchi MR, Rao MS, et al. A targeted neuroglial reporter line generated by homologous recombination in human embryonic stem cells. Stem Cells. 2009;27(8):1836–46.
35. Aizawa E, Hirabayashi Y, Iwanaga Y, Suzuki K, Sakurai K, Shimoji M, Aiba K, Wada T, Tooi N, Kawase E, et al. Efficient and accurate homologous recombination in hESCs and hiPSCs using helper-dependent adenoviral vectors. Mol Ther. 2012;20(2):424–31.
36. Zhang PW, Haidet-Phillips AM, Pham JT, Lee Y, Huo Y, Tienari PJ, Maragakis NJ, Sattler R, Rothstein JD. Generation of GFAP::GFP astrocyte reporter lines from human adult fibroblast-derived iPS cells using zinc-finger nuclease technology. Glia. 2016;64:63–75.
37. Davis RP, Ng ES, Costa M, Mossman AK, Sourris K, Elefanty AG, Stanley EG. Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak-like cells and enables isolation of primitive hematopoietic precursors. Blood. 2008;111(4):1876–84.
38. Elliott DA, Braam SR, Koutsis K, Ng ES, Jenny R, Lagerqvist EL, Biben C, Hatzistavrou T, Hirst CE, Yu QC, et al. NKX2-5(eGFP/w) hESCs for isolation of human cardiac progenitors and cardiomyocytes. Nat Methods. 2011;8(12):1037–40.
39. Den Hartogh SC, Schreurs C, Monshouwer-Kloots JJ, Davis RP, Elliott DA, Mummery CL, Passier R. Dual reporter MESP1 mCherry/w-NKX2-5 eGFP/w hESCs enable studying early human cardiac differentiation. Stem Cells. 2015;33(1):56–67.
40. Resnick MA. The repair of double-strand breaks in DNA: a model involving recombination. J Theor Biol. 1976;59(1):97–106.

41. Zou J, Maeder ML, Mali P, Pruett-Miller SM, Thibodeau-Beganny S, Chou BK, Chen G, Ye Z, Park IH, Daley GQ, et al. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell. 2009;5(1):97–110.
42. Cohen-Tannoudji M, Robine S, Choulika A, Pinto D, El Marjou F, Babinet C, Louvard D, Jaisser F. I-SceI-induced gene replacement at a natural locus in embryonic stem cells. Mol Cell Biol. 1998;18(3):1444–8.
43. Beerli RR, Barbas III CF. Engineering polydactyl zinc-finger transcription factors. Nat Biotechnol. 2002;20(2):135–41.
44. Segal DJ, Dreier B, Beerli RR, Barbas III CF. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. Proc Natl Acad Sci U S A. 1999;96(6):2758–63.
45. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, Ando D, Urnov FD, Galli C, Gregory PD, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol. 2007;25(11):1298–306.
46. Ptaszek LM, Cowan CA. New tools for genome modification in human embryonic stem cells. Cell Stem Cell. 2007;1(6):600–2.
47. Wu J, Kandavelou K, Chandrasegaran S. Custom-designed zinc finger nucleases: what is next? Cell Mol Life Sci. 2007;64(22):2933–44.
48. Bogdanove AJ, Voytas DF. TAL effectors: customizable proteins for DNA targeting. Science. 2011;333(6051):1843–6.
49. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011;29(8):731–4.
50. Harrison MM, Jenkins BV, O'Connor-Giles KM, Wildonger J. A CRISPR view of development. Genes Dev. 2014;28(17):1859–72.
51. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281–308.
52. Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. Cell Stem Cell. 2013;12(4):393–4.
53. Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, Erdin S, Cowan CA, Talkowski ME, Musunuru K. Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell. 2014;15(1):27–30.
54. Miki K, Endo K, Takahashi S, Funakoshi S, Takei I, Katayama S, Toyoda T, Kotaka M, Takaki T, Umeda M, et al. Efficient detection and purification of cell populations using synthetic MicroRNA switches. Cell Stem Cell. 2015;16(6):699–711.
55. Kamp TJ. Recognizing heart cells in a crowd. Nat Methods. 2011;8(12):1013–6.
56. Den Hartogh SC, Passier R. Fluorescent reporters in human pluripotent stem cells: contributions to cardiac differentiation and their applications in cardiac disease and toxicity. Stem Cells. 2016;34:13–26.
57. Kita-Matsuo H, Barcova M, Prigozhina N, Salomonis N, Wei K, Jacot JG, Nelson B, Spiering S, Haverslag R, Kim C, et al. Lentiviral vectors and protocols for creation of stable hESC lines for fluorescent tracking and drug resistance selection of cardiomyocytes. PLoS One. 2009;4(4), e5046.
58. Ovchinnikov DA, Hidalgo A, Yang SK, Zhang X, Hudson J, Mazzone SB, Chen C, Cooper-White JJ, Wolvetang EJ. Isolation of contractile cardiomyocytes from human pluripotent stem-cell-derived cardiomyogenic cultures using a human NCX1-EGFP reporter. Stem Cells Dev. 2015;24(1):11–20.
59. Xu XQ, Zweigerdt R, Soo SY, Ngoh ZX, Tham SC, Wang ST, Graichen R, Davidson B, Colman A, Sun W. Highly enriched cardiomyocytes from human embryonic stem cells. Cytotherapy. 2008;10(4):376–89.
60. Ross RS, Navankasattusas S, Harvey RP, Chien KR. An HF-1a/HF-1b/MEF-2 combinatorial element confers cardiac ventricular specificity and established an anterior-posterior gradient of expression. Development. 1996;122(6):1799–809.
61. Zhu WZ, Xie Y, Moyes KW, Gold JD, Askari B, Laflamme MA. Neuregulin/ErbB signaling regulates cardiac subtype specification in differentiating human embryonic stem cells. Circ Res. 2010;107(6):776–86.

Current Status of Genome Editing in Cardiovascular Medicine

Vittavat Termglinchan, Ioannis Karakikes, Timon Seeger, and Joseph C. Wu

Abstract During the past decades, numerous genetic mutations have been implicated in the pathogenesis of cardiovascular diseases (CVDs). With the launching of the Precision Medicine Initiative in January 2015, emerging technologies such as induced pluripotent stem cells (iPSCs) and genome editing are well positioned to provide powerful tools to correlate genotypes with phenotypes. These new technologies are helping to identify specific mutations associated with human CVDs. Patient-specific iPSC-derived cardiomyocytes (iPSC-CMs) offer an exciting experimental platform to model CVDs, study the molecular basis of cardiovascular biology, accelerate predictive drug toxicology tests, and advance potential regenerative therapies. By harnessing the power of genome engineering, scientists are uncovering the molecular mechanisms underlying the pathogenesis of CVDs, which will pave the way for the development of new personalized prediction, prevention, and treatment of diseases.

Keywords Genetic cardiovascular diseases • Cardiovascular disease modeling • Induced pluripotent stem cells • Genome engineering • Precision medicine

V. Termglinchan • I. Karakikes • T. Seeger
Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA 94305, USA

Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

J.C. Wu (✉)
Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA 94305, USA

Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, 265 Campus Drive, Room G1120, Stanford, CA 94305, USA
e-mail: joewu@stanford.edu

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_7

Abbreviations

AAV	Adeno-associated virus
APD	Action potential duration
ARVD/C	Arrhythmogenic right ventricular dysplasia/cardiomyopathy
BAV	Bicuspid aortic valve
BTHS	Barth syndrome
CHD	Coronary heart disease
CPVT	Catecholaminergic polymorphic ventricular tachycardia
CRISPR/Cas	Clustered regularly interspaced short palindromic repeat (CRISPR)/ Cas (CRISPR-associated)
CVD	Cardiovascular disease
DCM	Dilated cardiomyopathy
DSB	Double-strand break
EAD	Early afterdepolarization
FH	Familial hypercholesterolemia
GWAS	Genome-wide association studies
HCM	Hypertrophic cardiomyopathy
HDR	Homology-directed repair
iPSC	Induced pluripotent stem cell
iPSC-CM	Induced pluripotent stem cell-derived cardiomyocyte
iPSC-EC	Induced pluripotent stem cell-derived endothelial cell
iPSC-SMC	Induced pluripotent stem cell-derived smooth muscle cell
KCNH2	Potassium channel voltage-gated eag related subfamily H, member 2
KCNQ1	Potassium channel voltage-gated KQT-like subfamily Q, member 1
LDL-C	Low-density lipoprotein cholesterol
LDLR	Low-density lipoprotein receptor
LQTS	Long-QT syndrome
LVNC	Left ventricular non-compaction
N1	Notch1
NHEJ	Nonhomologous end-joining
PCSK9	Proprotein convertase subtilisin/kexin type 9
PLN	Phospholamban
RCM	Restrictive cardiomyopathy
SaCas9	*Staphylococcus aureus* Cas9
SpCas9	*Streptococcus pyogenes* Cas9
TALEN	Transcription activator-like effector nuclease
TAZ	Tafazzin
TTN	Titin
VUS	Variant of uncertain significance
ZFN	Zinc-finger nuclease

Cardiovascular Disease Genomics

Precision Medicine

Cardiovascular disease (CVD) is a major health problem that affects more than 85 million individuals in the United States alone [1]. One of the key goals in biomedical research currently is to identify the specific genes and variants associated with CVDs in humans. Such anticipated outcomes promise to transform human health by enabling more personalized prediction, prevention, and treatment of cardiovascular diseases on an individual level. This precision medicine approach is principally based on the ability to diagnose and stratify patients into different treatment groups by correlating a patient's genotype with the associated cellular phenotype, which will indicate how the genetic differences among individuals could influence their responses to therapies [2]. However, realizing the potential to treat individual patients requires the development of an accurate and cost-effective diagnosis system and reliable disease models.

The Genetics of Cardiovascular Disease

Human cardiovascular diseases include a wide range of disorders, including congenital heart diseases, cardiomyopathies, vasculature, and electrical conduction disorders. Recent advances have shown that genetics are significant in conferring risk for these disorders [3]. More than 1000 genetic mutations in more than 100 genes have been associated with inherited CVDs [3, 4], including dilated and hypertrophic cardiomyopathy (DCM and HCM, respectively) [5–7], arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D) [5, 8], long-QT syndromes (LQTS) [9, 10], aortic aneurysms [11], and hypercholesterolemia [12].

In 1975, Goldstein and colleagues identified a homozygous missense mutation in the low-density lipoprotein receptor (LDLR) gene in a patient with familial hypercholesterolemia (FH), representing the first demonstration of a causal genetic variant in Mendelian CVD [13–16]. Since then, many Mendelian forms (monogenic disorders) of CVD have been successfully identified by direct DNA sequencing or linkage analysis (Fig. 1) [4, 17]. However, this specific pattern of inheritance is rare and constitutes a minority of cases. Most of the common CVDs involve multiple genes, and their inheritance patterns can be variable and complex [18]. One of the main challenges in genetic research is to identify the genes that contribute to complex diseases. To this end, the publication of the first draft of the human genome in 2001 provided a valuable resource of detailed information about the structure, organization, and function of the nearly complete set of human genes [19, 20].

A decade later, the genome-wide association studies (GWAS), which examined genetic variants to determine the disease-causing variants between case and control subjects, have identified hundreds of loci associated with CVDs and traits [21]. Although

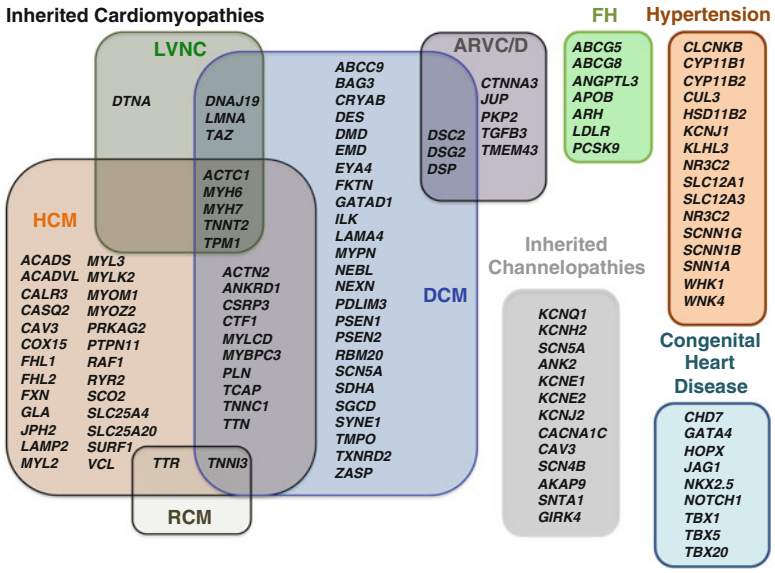

Fig. 1 The genetic basis of Mendelian cardiovascular diseases. Mutations in more than 100 genes have been associated with Mendelian cardiovascular diseases. Inherited cardiomyopathies, characterized by significant overlap, include hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), restrictive cardiomyopathy (RCM), left ventricular non-compaction (LVNC), and arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D). *FH* familial hypercholesterolemia

a plethora of newly discovered loci associated with cardiovascular risk factors and disease have been reported, the application of these findings to diagnosis, risk prediction, prevention and treatment of disease is still in its infancy and requires further research.

In the span of just a few years, rapid advances in next generation sequencing technology, either targeted or genome-wide, have identified and will continue to discover numerous novel genes associated with CVD. Targeted sequencing is now used to sequence candidate regions of the human genome. In a recent study, Wilson et al. demonstrated that cardiomyopathy-associated gene mutations can be identified with high fidelity using a high-throughput, clinical-grade next-generation targeted sequencing assay, providing a powerful new tool for CVD variant discovery [17]. Genome-wide DNA sequencing consists of whole-exome and whole-genome sequencing. In whole-exome sequencing, rare genetic variants of CVD can be identified by sequencing the protein-coding regions in large cohorts with a strong evidence of heritability. Because the majority of genomic content is constituted of noncoding regions, whole-genome sequencing is a comprehensive approach to identifying novel variants in both coding and noncoding regions [19, 20]. Data from the ENCODE project suggest that 37 % of the total human genome might have a function and is probably regulated in a tissue-specific manner [22]. Recent studies by Cordell et al. demonstrated that mutations in the noncoding genomic regions are strongly associated with multiple congenital heart diseases, including tetralogy

of Fallot [23, 24]. Therefore, as whole-genome sequencing becomes more widely utilized, more pathogenic variants associated with coding and noncoding RNAs will be uncovered.

Disease Modeling with Human Pluripotent Stem Cells

The molecular mechanisms underlying the pathogenesis of CVDs remain poorly understood despite tremendous advances in genetics. Disease models have been and will continue to provide important insights into the molecular basis of cardiovascular biology and disease. Transgenic animal models and heterologous cell systems have proven to be highly valuable in understanding of human CVDs [25]. However, considerable differences exist between animal models and human cells, and therefore human-based models are particularly important for cardiovascular research [26].

The recent discovery of the human iPSC technology [27–30], and improvements in the differentiation method of iPSCs into disease-relevant cell types such as cardiomyocytes (iPSC-CMs) [31–34], smooth muscle cells (iPSC-SMCs) [35, 36], and endothelial cells (iPSC-ECs) [37, 38], now provide an unprecedented opportunity for the generation of human patient-specific cell-based models for disease modeling, personalized drug screening, and regenerative approaches [39]. Indeed, significant progress has been made in iPSC-CM technology, which has been used to model monogenic diseases in vitro [40, 41]. Diseases such as LQT (LQTS1 [42–44], LQTS2 [45–47], LQTS3 [48], LQTS8/Timothy syndrome [49]), catecholaminergic polymorphic ventricular tachycardia (CPVT) [50–51], ARVC/D [52–53], HCM [54], and DCM [55–59] have been further elucidated using iPSC-CMs. In principle, this technology provides a means by which a patient's pathophysiology can be studied in vitro. However, the extent to which studies using patient-derived iPSCs will offer any advantage in understanding CVD pathogenesis is yet to be determined.

In addition, the interpretation of any phenotypes observed in a patient's iPSC-derived cells can only be understood via comparison with appropriate control cells. The iPSC-based disease models do not account for possible confounders of genetic background differences between patient iPSCs that might be responsible for the phenotypic differences. Even in studies where healthy siblings have been used as controls for disease patients, the phenotypic differences observed could be the result of natural variance in the genome, rather than in the putative disease-associated mutations [60]. Thus, the ideal comparison would be between cell lines that differ only in the genetic variant (i.e., isogenic cell lines). One way to ensure a better comparison would be to use isogenic wild-type control and mutated cell lines derived by site-specific genome editing from the same parental cell line (Fig. 2).

Site-specific genome-editing technology enables targeted double-strand breaks (DSBs) of the DNA at chromosomal loci of interest [61, 62]. DSBs subsequently recruit endogenous repair machinery for either nonhomologous end-joining (NHEJ) or homology-directed repair (HDR) pathways. The NHEJ

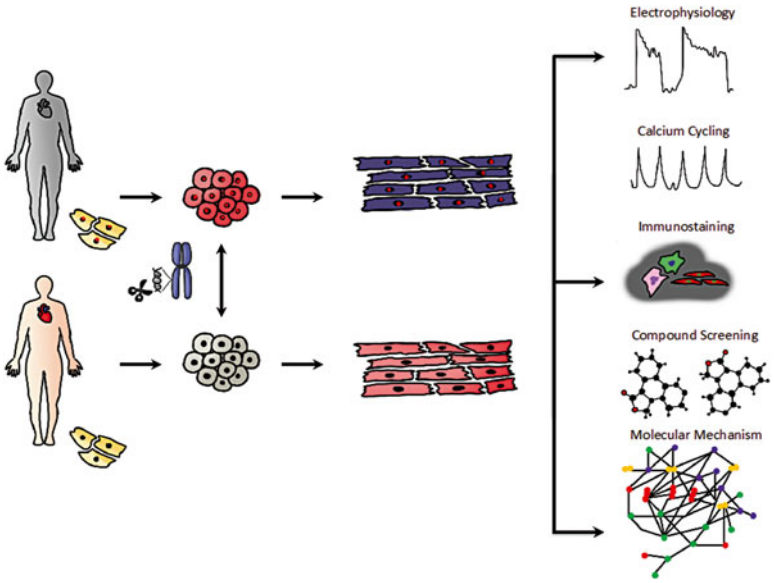

Fig. 2 Utilizing induced pluripotent stem cells (iPSC) and genome-editing technologies to model cardiovascular diseases. The genetic variability among human iPSCs can affect the outcome in modeling experiments. Site-specific genome engineering could in principle eliminate the variation arising from the iPSC line derivation and the genetic background. Isogenic iPSC lines (wild-type control and mutated cell lines) can be generated using the parental iPSC lines derived from healthy controls or patients. Generation of isogenic iPSC-CMs can be used to establish whether the observed in vitro phenotypes are the direct results of the disease-causing variant independently of the genetic background noise. Functional assays can be performed to identify disease-related molecular mechanisms. Additionally, this approach can be utilized in compound screening assays to find novel therapies and to determine the susceptibility of the genetic variant to drug-induced cardiac toxicity and arrhythmias. iPSCs induced pluripotent stem cells; iPSC-CMs induced pluripotent stem cell-derived cardiomyocytes

pathway generates random insertions or deletions at the site of DSBs, whereas HDR employs homologous donor DNA sequences from sister chromatids, homologous chromosomes, or exogenous DNA templates to produce precise genetic alteration. NHEJ and HDR are utilized for different aspects of genome engineering: NHEJ is specifically for gene inactivation, whereas HDR is for precise gene insertions, corrections, deletions, or base substitutions.

To date, four major classes of programmable nucleases, including meganucleases and their derivatives [63–66], zinc-finger nucleases (ZFNs) [67–76], transcription activator-like effector nucleases (TALENs) [77–85], and the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas (CRISPR-associated) (CRISPR/Cas) system [86–94] have been developed to enable site-specific genome engineering in a precise and predictable manner (Table 1). Indeed, genes have been inserted into specific loci, and gene mutations have been introduced or corrected in human iPSC-based cardiovascular disease models. An increasing number of studies utilize genome editing and iPSC technologies to not only study the biological mechanisms of genetic CVD but also to provide personalized therapies for these diseases (Table 2).

Table 1 Comparison of three classes of site-specific nucleases

	ZFNs	TALENs	CRISPR/Cas
DNA-binding determinant	Zinc-finger proteins	Transcription activator-like effectors	crRNA or sgRNA
Nucleases	*Fok*I	*Fok*I	Cas
Success rate[a]	Low (~20–30 %)	High (>90 %)	High (>90 %)
Average disruption rate[a,b]	Low (~10 %)	High (~20 %)	High (~20 %)
Length of target site	18–36 bp	30–40 bp	23 bp
Restriction in target site	G-rich	Start with T and end with A	End with NGG or NAG sequence
Off-target effect	High	Low	Variable
Size	~1 kb × 2	~3 kb × 2	3.3 kb (SaCas9) or 4.2 kb (SpCas9) + 0.1 kb (sgRNA)

CRISPR/*Cas* clustered regularly interspaced short palindromic repeat (CRISPR)/Cas (CRISPR-associated), *crRNA* CRISPR RNA, *sgRNA* single chain guide RNA, *SaCas9 Staphylococcus aureus* Cas9, *SpCas9 Streptococcus* pyogenes Cas9, *TALEN* transcription activator-like effector nuclease, *ZFN* zinc-finger nuclease

[a]The success rates are based on studies using HEK293 cells [61, 120–125]

[b]The average disruption rate is based on the frequency of nonhomologous end-joining at the nuclease target site

Table 2 Summary of major efforts using genome-editing technology to model and treat cardiovascular diseases

Disorder	Study	Gene	Platform	Findings
Disease modeling of genetic cardiovascular diseases				
BTHS	Wang et al. [97]	*TAZ* (c.517delG, c.328 T>C)	CRISPR/Cas9-mediated NHEJ and HDR (random insertion and gene mutation, respectively)	Immature cardiolipin content, abnormal mitochondrial functions, impaired sarcomere organization, and depressed contractile stress generation
DCM	Karakikes et al. [56]	*PLN* (p.R14del)	TALEN-mediated HDR (gene correction)	Calcium-handling abnormalities and abnormal *PLN* protein distribution
	Hinson et al. [57]	*TTN* (p.W976R, c.V6382fs, p.A22352fs, p.P22582fs, c.N22577fs, c.T33520fs)	CRISPR/Cas9-mediated HDR (gene correction, mutation)	Sarcomere insufficiency, impaired responses to mechanical stress and beta-adrenergic signaling with A-band mutations

(continued)

Table 2 (continued)

Disorder	Study	Gene	Platform	Findings
LQTS1	Wang et al. [42]	*KCNQ1* (p.R190Q, p.G269S, p.G345E)	ZFN-mediated HDR (gene insertion)	Prolonged APD and EADs
LQTS2	Wang et al. [42]	*KCNH2* (p.A614V)	ZFN-mediated HDR (gene insertion)	Prolonged APD and EADs
	Bellin et al. [100]	*KCNH2* (p.N996I)	Conventional homologous recombination	Reduced cell membrane *KCNH2* channels, decreased IKr, prolonged APD
BAV	Theodoris et al. [102]	*N1* (p.R1108X, p. H1505del)	TALEN-mediated HDR (gene correction)	Epigenetic dysregulation resulting in derepression of pro-osteogenic and pro-inflammatory gene networks
Personalized therapy of genetic cardiovascular diseases				
FH	Ding et al. [111]	*PCSK9*	CRISPR/ Cas9-mediated NHEJ (exon 1) delivered by adenovirus	Decrease plasma *PCSK9* levels and plasma LDL-C
	Ran et al. [118]	*PCSK9*	CRISPR/ Cas9-mediated NHEJ (exon 1 and 5) delivered by AAV	Decrease plasma *PCSK9* levels and total cholesterol level

AAV adeno-associated virus, *APD* action potential duration, *BAV* bicuspid aortic valve, *BTHS* Barth syndrome, *CRISPR*/*Cas* clustered regularly interspaced short palindromic repeat (CRISPR)/ Cas (CRISPR-associated), *DCM* dilated cardiomyopathy, *EAD* early afterdepolarization, *FH* familial hypercholesterolemia, *HDR* homology-directed repair, *KCNH2* potassium channel, voltage-gated eag-related subfamily H, member 2, *KCNQ1* potassium channel, voltage-gated KQT-like subfamily Q, member 1, *LDL-C* low-density lipoprotein cholesterol, *LQTS* long-QT syndrome, *NHEJ* nonhomologous end-joining, *N1* notch1, *PCSK9* proprotein convertase subtilisin/kexin type 9, *PLN* phospholamban, *TALEN* transcription activator-like effector nuclease, *TAZ* tafazzin, *TTN* titin, *ZFN* zinc-finger nuclease

Disease Modeling of Genetic Cardiovascular Disease

Inherited Cardiomyopathies

Exciting progress has been made in defining the etiology of inherited cardiomyopathies, including HCM, DCM, restrictive cardiomyopathy (RCM), left ventricular non-compaction cardiomyopathy (LVNC), and ARVC/D [40, 95]. To date,

numerous mutations in more than 50 genes that are associated with the pathogenesis of inherited cardiomyopathies have been discovered [17]. Although molecular analysis efforts have revealed important insights regarding the role of genetics in cardiomyopathies, the underlying molecular mechanisms of HCM, DCM, RCM, LVNC, and ARVC/D remain unclear.

In recent years, the iPSC-CM technology has been used to model inherited cardiomyopathies [5, 41]. However, one of the major limitations still remains the variability resulting from the genetic background differences between iPSC lines. In a study by Wang et al., human iPSC-CMs were generated from two patients with Barth syndrome (BTHS), an inherited X-linked cardiac and skeletal mitochondrial myopathy caused by mutation of the gene encoding for tafazzin (*TAZ*) [96, 97]. The study used CRISPR/Cas9-mediated NHEJ to mutate *TAZ* and demonstrated that the mutated isogenic iPSC-CMs exhibited similar phenotypes as BTHS patient-specific iPSC-CMs, including immature cardiolipin content, abnormal mitochondrial functions, impaired sarcomere organization, and depressed contractile stress generation.

Genome-editing technology has also been utilized to study the pathogenesis of familial DCM. In a study by Karakikes et al., the p.R14del mutation in the coding region of the phospholamban (*PLN*) gene was corrected by TALEN-mediated homology-directed repair (HDR) in patient-specific iPSCs [56]. After differentiation into cardiomyocytes, the DCM phenotype was ameliorated in TALEN-corrected iPSC-CMs when compared to the isogenic *PLN* mutated cells, including alleviation of Ca^{2+}-handling abnormalities, electrical instability, and abnormal cytoplasmic distribution of the *PLN* protein.

Most recently, Hinson et al. utilized iPSCs and genome-editing technologies to evaluate the pathogenicity of titin (*TTN*) gene variants [57]. Their study used CRISPR/Cas9-mediated homologous recombination to introduce and correct either missense or frameshift mutations in several loci of the *TTN* gene, including four mutations affecting the A-band and two mutations impacting the I-band. By combining functional analyses with RNA sequencing of isogenic lines, they demonstrated that mutations in the A-band domain of the *TTN* cause DCM, whereas truncations in the I-band are better tolerated. This study also showed that the pathogenesis of *TTN*-induced DCM is associated with sarcomere insufficiency, as well as impaired responses to mechanical stress, and abnormal beta-adrenergic signaling.

Inherited Channelopathies

Long-QT syndrome (LQTS) is an inherited or acquired cardiac arrhythmic disease, predisposing patients to life-threatening ventricular arrhythmias and sudden cardiac death [9, 10]. Mutations in 13 genes have been implicated in the pathogenesis of familial LQTS [98]. The potassium channels, voltage-gated KQT-like subfamily Q, member 1 (*KCNQ1*; LQTS1), and voltage-gated eag-related subfamily H, member

2 (*KCNH2*; LQTS2) are the most common mutated genes associated with LQTS. Because both *KCNQ1* and *KCNH2* function as homotetramers, the mutated monomer displays a dominant-negative effect by impairing the tetramerization of wild-type monomers [42].

In the past 5 years, LQTS1 [42–44], LQTS2 [45–47], LQTS3 [48], and LQTS8/Timothy syndrome [49] have been modeled by iPSC-CMs [9, 99]. However, this approach can still be somewhat limited because it is difficult to obtain patient samples with the desired genetic variants. A study by Wang et al. represented an exciting first step in producing human cardiomyocytes that recapitulated LQTS by inserting the mutated genes in the safe-harbor locus (*AAVS1*/*PPP1R12C*) [42]. Their study utilized ZFNs to insert an expression cassette encoding a pathogenic variant of *KCNQ1* and *KCNH2* into wild-type iPSCs. The *KCNQ1*-mutated and *KCNH2*-mutated iPSC-CMs showed prolonged action potential duration (APD) and calcium-handling abnormalities when compared to the isogenic control iPSC-CMs. This study demonstrated an alternative approach to using actual patient samples and represents a novel way to study genetic variants that are known to display dominant negative effects.

Another study by Bellin et al. utilized a conventional HDR strategy (without using site-specific nucleases) to generate isogenic mutated and wild-type lines of a heterozygous missense *KCNH2* p.N996I mutation [100]. Correction of the mutation restored the electrical current conducted by the HERG channel (IKr) and the action potential duration in iPSC-CMs. As expected, introduction of the same genetic mutation reduced the IKr and prolonged the action potential duration in iPSC-CMs. Their study demonstrated that the isogenic mutated iPSC-CMs expressed fewer *KCNH2* channels at the cell membrane than the isogenic wild-type iPSC-CMs. Further treatment with the proteasome inhibitors, lactacystin and leupeptin, increased the protein levels of *KCNH2* on the cell membrane in the mutated iPSC-CMs, which may suggest a role of proteasomes in the pathogenesis of LQTS2.

Inherited Valvulopathies

In the past decade, congenital defects of the aortic valve known as bicuspid aortic valve (BAV) have been associated with genetic variants of a membrane-bound transcription factor, NOTCH1 (*N1*) [101]. BAV occurs in 1–2 % of the population and involves the formation of two valve leaflets instead of the normal three leaflets. Although the mechanism remains largely unknown, BAV is a major risk factor for early aortic valve calcification, a condition that impairs the opening of the valve and is responsible for more than 100,000 valve transplants annually in the United States.

A study by Theodoris et al. recruited two families carrying a *N1* heterozygous nonsense mutation, which is suspected to cause congenital BAV [102]. Their study utilized TALEN-mediated HDR to correct the *N1* mutation in patient-specific iPSC lines. Comparing the isogenic mutated and control human iPSC-ECs, the *N1* mRNA levels were found to be reduced by 30–40 % in the isogenic mutated iPSC-ECs, which indicates that the *N1* mutation displays a haploinsufficient effect on the

pathogenesis of BAV. After exposing the isogenic iPSC-ECs to shear stress, the *N1* isogenic mutated cells demonstrated epigenetic dysregulation, resulting in derepression of latent pro-osteogenic and pro-inflammatory gene networks.

Personalized Therapy of Genetic Cardiovascular Diseases

Challenges in Therapeutic Genome Engineering

Although different cell types have distinct abilities to repair DSBs, the phase of the cell cycle primarily governs the choice of whether the NHEJ or HDR pathways are utilized. NHEJ dominates DNA repair during the G_1-, S-, and G_2-phases, whereas HDR is limited to the late S- and G_2-phases [103, 104]. This difference in cellular activity makes it more challenging to treat diseases that require HDR-mediated DSB repair (gene correction or gene insertion) than those requiring NHEJ-mediated repair (gene inactivation). Several studies have demonstrated that chemical or genetic interruption of the NHEJ pathway can favor HDR; however, such manipulations can be difficult to employ and are harmful to cells [105, 106].

The potential use of genome-editing technology in cardiovascular therapy can be divided into two approaches: those carried out in vitro or those in vivo (Fig. 3a, b, respectively). With established in vitro approaches, the editing pro-

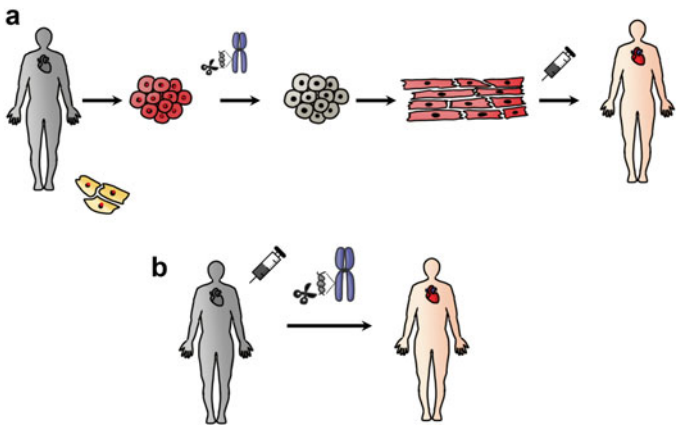

Fig. 3 Personalized therapy of genetic cardiovascular disease (CVD). The potential use of genome-editing technology in cardiovascular therapy can be divided in two approaches: those carried out in vitro or those in vivo. (**a**) With established in vitro approaches, the editing process is achieved via inpatient-specific iPSC lines in culture, and these iPSC lines are subsequently differentiated into the human cell type of interest by using established differentiation protocols. These edited cells can be delivered to patients to treat specific CVDs. (**b**) In vivo genome-editing therapy is achieved by delivering programmable nucleases to patients to correct or disrupt the mutations of interest

cess is achieved in human iPSC lines in culture, and these iPSC lines are subsequently differentiated into the human cell type of interest, such as iPSC-CMs or iPSC-ECs, by using established differentiation protocols [31–34, 37, 38]. These corrected cells can be delivered to patients to treat specific CVDs. By contrast, in vivo genome-editing therapy is still very much a work in progress. For instance, one limitation is that adult cardiomyocytes are arrested in the G_0-phase, in which the HDR mechanism is inactive and NHEJ is very limited. Furthermore, adult cardiomyocytes have a very low rate of replication [107], so that to achieve therapeutic effects, the efficiency of modification must be quite high. Nonetheless, in vivo genome-editing therapy is technically more feasible in certain cell types, including hepatocytes and satellite cells, mainly because these types of cells replicate better and theoretically could outcompete the native diseased cells [108–110].

Current Targets for In Vivo Genome-Editing Therapy

The concentration of low-density lipoprotein cholesterol (LDL-C) in the blood is among the most established causal risk factors for coronary heart disease (CHD) [111]. Pharmacological agents that reduce LDL-C levels, namely statins, are currently the most effective means of reducing this coronary heart disease risk. However, a large proportion of patients are intolerant to statin therapy. Proprotein convertase subtilisin/kexin type 9 (*PCSK9*) has been identified as the cause of autosomal dominant FH [112]. *PCSK9* is specifically expressed in and secreted from the liver and functions as an antagonist to the LDLR. Therefore, *PCSK9* has now emerged as a promising therapeutic target for the prevention of CHD.

Studies have shown that individuals with loss-of-function mutations in *PCSK9* experienced a significant reduction of LDL-C levels and consequently CHD risk [113–116]. As might be expected, gain-of-function mutations elevate LDL-C levels, leading to early-onset CHD in patients diagnosed with FH [112]. In 2015, a *PCSK9*-targeting monoclonal antibody (alirocumab) was approved by the U.S. Food and Drug Administration for the treatment of FH [117]. Although this antibody has been shown to be effective for the treatment of FH, its effects on LDL-C are transient. Patients must receive injections every few weeks to maintain the desired level of *PCSK9*-targeting monoclonal antibody.

A study by Ding et al. utilized CRISPR/Cas9-mediated NHEJ to permanently disrupt the mouse *PCSK9* gene in vivo [111]. They demonstrated that after the administration of adenovirus to express a CRISPR/Cas9 targeting *PCSK9* in mice, at least 50 % of the *PCSK9* alleles in the liver were altered; this resulted in a significant decrease in plasma *PCSK9* levels, as well as an increase in hepatic LDLR levels and a reduction of plasma LDL-C level by 35–40 %. Their study is the first to demonstrate the potential of genome engineering in vivo for the prevention of CVD.

More recently, a study by Ran et al. demonstrated that Cas9 from *Staphylococcus aureus* (SaCas9) can alter the genome with an efficiency similar to that of Cas9 from *Streptococcus pyogenes* (SpCas9), while being encoded by a gene that is more than 1 kilobase (kb) shorter [118]. The smaller SaCas9 (~3.3 kb) enabled in vivo genome engineering using adeno-associated virus (AAV) that would otherwise be prohibited by the AAV's restrictive cargo size (~4.5 kb) [119]; the study then utilized CRISPR/SaCas9 to target the *PCSK9* gene in the mouse liver. As in the previous study [111], more than 40 % of gene modification was observed, accompanied by significant reductions in serum *PCSK9* and total cholesterol levels. Assessment of off-target effects by targeted deep sequencing did not show any indel formations. This study suggested that in vivo genome editing using the novel CRISPR/SaCas9 has the potential to be both highly efficient and specific.

Conclusions and Future Perspectives

With the launching of the Precision Medicine Initiative, rapidly emerging technologies such as iPSCs and genome editing are well positioned to provide powerful tools for studying genotype–phenotype associations, for predicting the cardiovascular risks of individual patients and their responses to therapies [2]. The iPSC technology is revolutionary and continues to evolve. As it becomes easier to edit mutations in iPSCs, it will become feasible to test genetic variants of uncertain significance (VUS), and to assess the importance of genetic modifiers on disease penetrance [9] (Fig. 4). The genome-editing technology presents a novel and rapidly advancing technology with exciting applications. However, significant challenges remain, including enhancing specificity and minimizing off-target effects, increasing efficiency, and improving the selection of targeted sites and delivery methods, and especially for in vivo genome engineering. Further refinements are needed to fully exploit the potential of genome editing to be a vital tool of future precision medicine treatment for CVD.

We envision that the use of genome engineering to generate human cell-based disease models will become a standard approach in the laboratory, allowing researchers to decipher the molecular mechanisms of genetic variants and unlock the secret of CVDs. Nevertheless, many obstacles remain unresolved at this point. Population-based data sets will be necessary to identify novel genetic variants that are contributors to CVDs. Bioinformatics will be an important tool to determine the casual relationship between genotypes and phenotypes, as most of the diseases in question will be polygenic. Finally, in the genetic diagnosis aspect, research should focus on improving the accuracy, flexibility, turnaround, and cost of the next-generation sequencing.

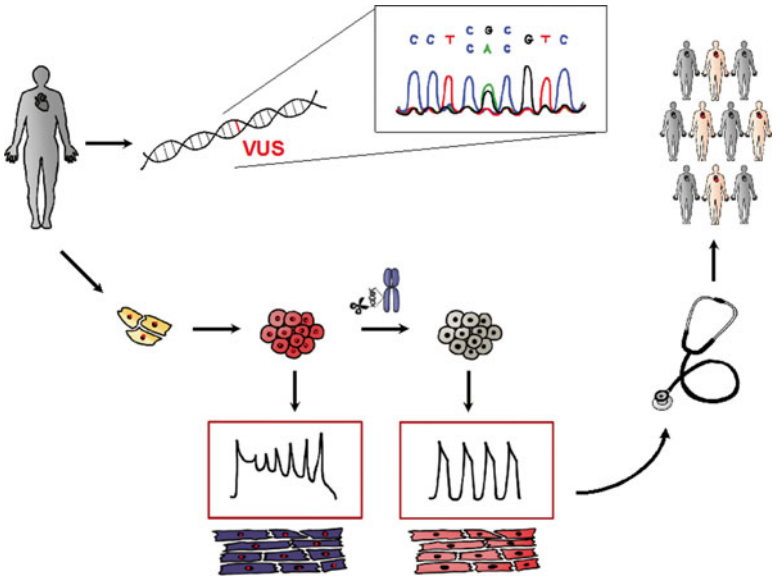

Fig. 4 Future perspectives of genome-editing technologies in cardiovascular research. Clinical genetic testing attempts to identify a rare variant in genes associated with CVDs. When a known pathogenic variant is identified, this may guide the clinical diagnosis, and genetic screening can be offered to family members to identify those at risk for developing the disease. When a variant of uncertain significance (VUS) is identified, a genome-corrected iPSCs can be developed, and functional assays can be done on the isogenic iPSC-CMs to detect the phenotypic abnormalities between isogenic wild-type control, and mutated iPSC-CMs; this may help to reclassify the VUS as possibly disease causing and place it in a similar category as a pathogenic variant. This technique can also be utilized to identify novel genetic modifiers or variants that may be associated with variability in drug responses

Acknowledgments We gratefully acknowledge Joseph Gold, Jade Chao, and Blake Wu for critical reading, and Varachaya Khwanjaipanich for preparing the illustrations. This work was supported from the National Institutes of Health (NIH) R01 HL113006, R01 HL123968, R01 HL125627, and R01 HL128170, and American Heart Association Established Investigator Award, Burroughs Wellcome Fund (J.C.W.), NIH K99 HL104002, AHA 15BGIA22730027, and Stanford CVI Seed Grant (I.K.), and German Research Foundation (T.S.).

References

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. Circulation. 2015;131(4):e29–39322.
2. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015;372(9):793–5.
3. Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. Cell. 2012;148(6):1242–57.

4. O'Donnell CJ, Nabel EG. Genomics of cardiovascular disease. N Engl J Med. 2011;365(22):2098–109.
5. Karakikes I, Termglinchan V, Wu JC. Human-induced pluripotent stem cell models of inherited cardiomyopathies. Curr Opin Cardiol. 2014;29(3):214–9.
6. Wigle ED, Rakowski H, Kimball BP, Williams WG. Hypertrophic cardiomyopathy. Clinical spectrum and treatment. Circulation. 1995;92(7):1680–92.
7. Dec GW, Fuster V. Idiopathic dilated cardiomyopathy. N Engl J Med. 1994;331(23):1564–75.
8. Corrado D, Thiene G. Arrhythmogenic right ventricular cardiomyopathy/dysplasia: clinical impact of molecular genetic studies. Circulation. 2006;113(13):1634–7.
9. Sallam K, Li Y, Sager PT, Houser SR, Wu JC. Finding the rhythm of sudden cardiac death: new opportunities using induced pluripotent stem cell-derived cardiomyocytes. Circ Res. 2015;116(12):1989–2004.
10. Priori SG, Schwartz PJ, Napolitano C, Bloise R, Ronchetti E, Grillo M, et al. Risk stratification in the long-QT syndrome. N Engl J Med. 2003;348(19):1866–74.
11. Loeys BL, Schwarze U, Holm T, Callewaert BL, Thomas GH, Pannu H, et al. Aneurysm syndromes caused by mutations in the TGF-beta receptor. N Engl J Med. 2006;355(8):788–98.
12. Hobbs HH, Brown MS, Russell DW, Davignon J, Goldstein JL. Deletion in the gene for the low-density-lipoprotein receptor in a majority of French Canadians with familial hypercholesterolemia. N Engl J Med. 1987;317(12):734–7.
13. Goldstein JL, Brown MS. The LDL receptor. Arterioscler Thromb Vasc Biol. 2009;29(4):431–8.
14. Brown MS, Goldstein JL. Lipoprotein receptors and genetic control of cholesterol metabolism in cultured human cells. Naturwissenschaften. 1975;62(8):385–9.
15. Goldstein JL, Dana SE, Brunschede GY, Brown MS. Genetic heterogeneity in familial hypercholesterolemia: evidence for two different mutations affecting functions of low-density lipoprotein receptor. Proc Natl Acad Sci U S A. 1975;72(3):1092–6.
16. Sudhof TC, Goldstein JL, Brown MS, Russell DW. The LDL receptor gene: a mosaic of exons shared with different proteins. Science. 1985;228(4701):815–22.
17. Wilson KD, Shen P, Fung E, Karakikes I, Zhang A, InanlooRahatloo K, et al. A rapid, high-quality, cost-effective, comprehensive and expandable targeted next-generation sequencing assay for inherited heart diseases. Circ Res. 2015;117(7):603–11.
18. Lee DS, Pencina MJ, Benjamin EJ, Wang TJ, Levy D, O'Donnell CJ, et al. Association of parental heart failure with risk of heart failure in offspring. N Engl J Med. 2006;355(2):138–47.
19. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001;291(5507):1304–51.
20. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860–921.
21. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet. 2003;33(2):177–82.
22. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57–74.
23. Cordell HJ, Topf A, Mamasoula C, Postma AV, Bentham J, Zelenika D, et al. Genome-wide association study identifies loci on 12q24 and 13q32 associated with tetralogy of Fallot. Hum Mol Genet. 2013;22(7):1473–81.
24. Cordell HJ, Bentham J, Topf A, Zelenika D, Heath S, Mamasoula C, et al. Genome-wide association study of multiple congenital heart disease phenotypes identifies a susceptibility locus for atrial septal defect at chromosome 4p16. Nat Genet. 2013;45(7):822–4.
25. Houser SR, Margulies KB, Murphy AM, Spinale FG, Francis GS, Prabhu SD, et al. Animal models of heart failure: a scientific statement from the American Heart Association. Circ Res. 2012;111(1):131–50.

26. Dixon JA, Spinale FG. Large animal models of heart failure: a critical link in the translation of basic science to clinical practice. Circ Heart Fail. 2009;2(3):262–71.
27. Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol. 2008;26(1):101–6.
28. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007;131(5):861–72.
29. Takahashi K, Okita K, Nakagawa M, Yamanaka S. Induction of pluripotent stem cells from fibroblast cultures. Nat Protoc. 2007;2(12):3081–9.
30. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;318(5858):1917–20.
31. Burridge PW, Matsa E, Shukla P, Lin ZC, Churko JM, Ebert AD, et al. Chemically defined generation of human cardiomyocytes. Nat Methods. 2014;11(8):855–60.
32. Burridge PW, Keller G, Gold JD, Wu JC. Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming. Cell Stem Cell. 2012;10(1):16–28.
33. Karakikes I, Senyei GD, Hansen J, Kong CW, Azeloglu EU, Stillitano F, et al. Small molecule-mediated directed differentiation of human embryonic stem cells toward ventricular cardiomyocytes. Stem Cells Transl Med. 2014;3(1):18–31.
34. Lian X, Bao X, Zilberter M, Westman M, Fisahn A, Hsiao C, et al. Chemically defined, albumin-free human cardiomyocyte generation. Nat Methods. 2015;12(7):595–6.
35. Bajpai VK, Mistriotis P, Loh YH, Daley GQ, Andreadis ST. Functional vascular smooth muscle cells derived from human induced pluripotent stem cells via mesenchymal stem cell intermediates. Cardiovasc Res. 2012;96(3):391–400.
36. Dash BC, Jiang Z, Suh C, Qyang Y. Induced pluripotent stem cell-derived vascular smooth muscle cells: methods and application. Biochem J. 2015;465(2):185–94.
37. Adams WJ, Zhang Y, Cloutier J, Kuchimanchi P, Newton G, Sehrawat S, et al. Functional vascular endothelium derived from human induced pluripotent stem cells. Stem Cell Rep. 2013;1(2):105–13.
38. Hu S, Zhao MT, Jahanbani F, Shao NY, Lee WH, Chen H, et al. Effects of cellular origin on differentiation of human induced pluripotent stem cell–derived endothelial cells. JCI Insight. 2016;1(8):e85558.
39. Matsa E, Sallam K, Wu JC. Cardiac stem cell biology: glimpse of the past, present, and future. Circ Res. 2014;114(1):21–7.
40. Karakikes I, Ameen M, Termglinchan V, Wu JC. Human induced pluripotent stem cell-derived cardiomyocytes: insights into molecular, cellular, and functional phenotypes. Circ Res. 2015;117(1):80–8.
41. Matsa E, Burridge PW, Wu JC. Human stem cells for modeling heart disease and for drug discovery. Sci Transl Med. 2014;6(239):239ps6.
42. Wang Y, Liang P, Lan F, Wu H, Lisowski L, Gu M, et al. Genome editing of isogenic human induced pluripotent stem cells recapitulates long QT phenotype for drug testing. J Am Coll Cardiol. 2014;64(5):451–9.
43. Egashira T, Yuasa S, Suzuki T, Aizawa Y, Yamakawa H, Matsuhashi T, et al. Disease characterization using LQTS-specific induced pluripotent stem cells. Cardiovasc Res. 2012;95(4):419–29.
44. Moretti A, Bellin M, Welling A, Jung CB, Lam JT, Bott-Flugel L, et al. Patient-specific induced pluripotent stem-cell models for long-QT syndrome. N Engl J Med. 2010;363(15):1397–409.
45. Matsa E, Rajamohan D, Dick E, Young L, Mellor I, Staniforth A, et al. Drug evaluation in cardiomyocytes derived from human induced pluripotent stem cells carrying a long QT syndrome type 2 mutation. Eur Heart J. 2011;32(8):952–62.

46. Lahti AL, Kujala VJ, Chapman H, Koivisto AP, Pekkanen-Mattila M, Kerkela E, et al. Model for long QT syndrome type 2 using human iPS cells demonstrates arrhythmogenic characteristics in cell culture. Dis Model Mech. 2012;5(2):220–30.
47. Itzhaki I, Maizels L, Huber I, Zwi-Dantsis L, Caspi O, Winterstern A, et al. Modelling the long QT syndrome with induced pluripotent stem cells. Nature. 2011;471(7337):225–9.
48. Davis RP, Casini S, van den Berg CW, Hoekstra M, Remme CA, Dambrot C, et al. Cardiomyocytes derived from pluripotent stem cells recapitulate electrophysiological characteristics of an overlap syndrome of cardiac sodium channel disease. Circulation. 2012;125(25):3079–91.
49. Yazawa M, Hsueh B, Jia X, Pasca AM, Bernstein JA, Hallmayer J, et al. Using induced pluripotent stem cells to investigate cardiac phenotypes in Timothy syndrome. Nature. 2011;471(7337):230–4.
50. Jung CB, Moretti A, Mederos y Schnitzler M, Iop L, Storch U, Bellin M, et al. Dantrolene rescues arrhythmogenic RYR2 defect in a patient-specific stem cell model of catecholaminergic polymorphic ventricular tachycardia. EMBO Mol Med. 2012;4(3):180–91.
51. Novak A, Barad L, Zeevi-Levin N, Shick R, Shtrichman R, Lorber A, et al. Cardiomyocytes generated from CPVTD307H patients are arrhythmogenic in response to beta-adrenergic stimulation. J Cell Mol Med. 2012;16(3):468–82.
52. Asimaki A, Kapoor S, Plovie E, Arndt AK, Adams E, Liu ZZ, et al. Identification of a new modulator of the intercalated disc in a zebrafish model of arrhythmogenic cardiomyopathy. Sci Transl Med. 2014;6(240):240ra74.
53. Kim C, Wong J, Wen J, Wang S, Wang C, Spiering S, et al. Studying arrhythmogenic right ventricular dysplasia with patient-specific iPSCs. Nature. 2013;494(7435):105–10.
54. Lan F, Lee AS, Liang P, Sanchez-Freire V, Nguyen PK, Wang L, et al. Abnormal calcium handling properties underlie familial hypertrophic cardiomyopathy pathology in patient-specific induced pluripotent stem cells. Cell Stem Cell. 2013;12(1):101–13.
55. Sun N, Yazawa M, Liu J, Han L, Sanchez-Freire V, Abilez OJ, et al. Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy. Sci Transl Med. 2012;4(130):130ra47.
56. Karakikes I, Stillitano F, Nonnenmacher M, Tzimas C, Sanoudou D, Termglinchan V, et al. Correction of human phospholamban R14del mutation associated with cardiomyopathy using targeted nucleases and combination therapy. Nat Commun. 2015;6:6955.
57. Hinson JT, Chopra A, Nafissi N, Polacheck WJ, Benson CC, Swist S, et al. Heart disease. Titin mutations in iPS cells define sarcomere insufficiency as a cause of dilated cardiomyopathy. Science. 2015;349(6251):982–6.
58. Wu H, Lee J, Vincent LG, Wang Q, Gu M, Lan F, et al. Epigenetic regulation of phosphodiesterases 2A and 3A underlies compromised β -adrenergic signaling in an iPSC model of dilated cardiomyopathy. Cell Stem Cell. 2015;17(1):89–100.
59. Davis J, Davis LC, Correll RN, Makarewich CA, Schwanekamp JA, Moussavi-Harami F, et al. A tension-based model distinguishes hypertrophic versus dilated cardiomyopathy. Cell. 2016;165(5):1147–59.
60. Musunuru K. Genome editing of human pluripotent stem cells to generate human cellular disease models. Dis Model Mech. 2013;6(4):896–904.
61. Kim H, Kim JS. A guide to genome engineering with programmable nucleases. Nat Rev Genet. 2014;15(5):321–34.
62. Rouet P, Smih F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol. 1994;14(12):8096–106.
63. Boissel S, Jarjour J, Astrakhan A, Adey A, Gouble A, Duchateau P, et al. megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. Nucleic Acids Res. 2014;42(4):2591–601.
64. Smith J, Grizot S, Arnould S, Duclert A, Epinat JC, Chames P, et al. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. Nucleic Acids Res. 2006;34(22), e149.

65. Thierry A, Dujon B. Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes. Nucleic Acids Res. 1992;20(21):5625–31.
66. Thierry A, Perrin A, Boyer J, Fairhead C, Dujon B, Frey B, et al. Cleavage of yeast and bacteriophage T7 genomes at a single site using the rare cutter endonuclease I-Sce I. Nucleic Acids Res. 1991;19(1):189–90.
67. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 2010;11(9):636–46.
68. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature. 2005;435(7042):646–51.
69. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. Science. 2003;300(5620):764.
70. Bibikova M, Golic M, Golic KG, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. Genetics. 2002;161(3):1169–75.
71. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A. 1996;93(3):1156–60.
72. Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol. 2007;25(7):778–85.
73. Miller J, McLachlan AD, Klug A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. 1985;4(6):1609–14.
74. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. Science. 2003;300(5620):763.
75. Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res. 2000;28(17):3361–9.
76. Wolfe SA, Nekludova L, Pabo CO. DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct. 2000;29:183–212.
77. Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, et al. A TALEN genome-editing system for generating human stem cell-based disease models. Cell Stem Cell. 2013;12(2):238–51.
78. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011;29(8):731–4.
79. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011;29(2):143–8.
80. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol. 2011;29(2):149–53.
81. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009;326(5959):1509–12.
82. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, et al. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 2010;186(2):757–61.
83. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res. 2011;39(1):359–72.
84. Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu JK. 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. 2011;108(6):2623–8.
85. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009;326(5959):1501.
86. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819–23.

87. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281–308.
88. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315(5819): 1709–12.
89. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005;151(Pt 8):2551–61.
90. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471(7340):602–7.
91. Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468(7320):67–71.
92. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337(6096):816–21.
93. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823–6.
94. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. Nucleic Acids Res. 2011;39(21):9275–82.
95. Cahill TJ, Ashrafian H, Watkins H. Genetic cardiomyopathies causing heart failure. Circ Res. 2013;113(6):660–75.
96. Bione S, D'Adamo P, Maestrini E, Gedeon AK, Bolhuis PA, Toniolo D. A novel X-linked gene, G4.5. is responsible for Barth syndrome. Nat Genet. 1996;12(4):385–9.
97. Wang G, McCain ML, Yang L, He A, Pasqualini FS, Agarwal A, et al. Modeling the mitochondrial cardiomyopathy of Barth syndrome with induced pluripotent stem cell and heart-on-chip technologies. Nat Med. 2014;20(6):616–23.
98. Schwartz PJ, Crotti L, Insolia R. Long-QT syndrome: from genetics to management. Circ Arrhythm Electrophysiol. 2012;5(4):868–77.
99. Sinnecker D, Goedel A, Laugwitz KL, Moretti A. Induced pluripotent stem cell-derived cardiomyocytes: a versatile tool for arrhythmia research. Circ Res. 2013;112(6):961–8.
100. Bellin M, Casini S, Davis RP, D'Aniello C, Haas J, Ward-van Oostwaard D, et al. Isogenic human pluripotent stem cell pairs reveal the role of a KCNH2 mutation in long-QT syndrome. EMBO J. 2013;32(24):3161–75.
101. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, et al. Mutations in NOTCH1 cause aortic valve disease. Nature. 2005;437(7056):270–4.
102. Theodoris CV, Li M, White MP, Liu L, He D, Pollard KS, et al. Human disease modeling reveals integrated transcriptional and epigenetic mechanisms of NOTCH1 haploinsufficiency. Cell. 2015;160(6):1072–86.
103. Heyer WD, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. Annu Rev Genet. 2010;44:113–39.
104. Rothkamm K, Kruger I, Thompson LH, Lobrich M. Pathways of DNA double-strand break repair during the mammalian cell cycle. Mol Cell Biol. 2003;23(16):5706–15.
105. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol. 2015;33(5):538–42.
106. Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. Cell Res. 2008;18(1):134–47.
107. Bergmann O, Bhardwaj RD, Bernard S, Zdunek S, Barnabe-Heider F, Walsh S, et al. Evidence for cardiomyocyte renewal in humans. Science. 2009;324(5923):98–102.

108. Cox DB, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. Nat Med. 2015;21(2):121–31.
109. Long C, McAnally JR, Shelton JM, Mireault AA, Bassel-Duby R, Olson EN. Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science. 2014;345(6201):1184–8.
110. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, et al. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat Biotechnol. 2014;32(6):551–3.
111. Ding Q, Strong A, Patel KM, Ng SL, Gosis BS, Regan SN, et al. Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. Circ Res. 2014;115(5):488–92.
112. Abifadel M, Varret M, Rabes JP, Allard D, Ouguerram K, Devillers M, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet. 2003;34(2):154–6.
113. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. Nat Genet. 2005;37(2):161–5.
114. Cohen JC, Boerwinkle E, Mosley Jr TH, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. 2006;354(12):1264–72.
115. Hooper AJ, Marais AD, Tanyanyiwa DM, Burnett JR. The C679X mutation in PCSK9 is present and lowers blood cholesterol in a Southern African population. Atherosclerosis. 2007;193(2):445–8.
116. Zhao Z, Tuakli-Wosornu Y, Lagace TA, Kinch L, Grishin NV, Horton JD, et al. Molecular characterization of loss-of-function mutations in PCSK9 and identification of a compound heterozygote. Am J Hum Genet. 2006;79(3):514–23.
117. Robinson JG, Farnier M, Krempf M, Bergeron J, Luc G, Averna M, et al. Efficacy and safety of alirocumab in reducing lipids and cardiovascular events. N Engl J Med. 2015;372(16):1489–99.
118. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. Nature. 2015;520(7546):186–91.
119. Senis E, Fatouros C, Grosse S, Wiedtke E, Niopek D, Mueller AK, et al. CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. Biotechnol J. 2014;9(11):1402–12.
120. Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol. 2013;31(3):230–2.
121. Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, et al. A library of TAL effector nucleases spanning the human genome. Nat Biotechnol. 2013;31(3):251–8.
122. Kim JS, Lee HJ, Carroll D. Genome editing with modularly assembled zinc-finger nucleases. Nat Methods. 2010;7(2):91. author reply 92.
123. Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res. 2009;19(7):1279–88.
124. Kim Y, Kweon J, Kim JS. TALENs and ZFNs are associated with different mutation signatures. Nat Methods. 2013;10(3):185.
125. Kim S, Lee MJ, Kim H, Kang M, Kim JS. Preassembled zinc-finger arrays for rapid construction of ZFNs. Nat Methods. 2011;8(1):7.

Using CRISPR-Cas9 Genome Editing to Enhance Cell Based Therapies for the Treatment of Diabetes Mellitus

Nicole A.J. Krentz and Francis C. Lynn

Abstract Human pluripotent stem cells (hPSCs) have the ability to form all cell types of the body, making them an excellent potential source of insulin-producing pancreatic β-cells for diabetes treatment. To generate these cells in vitro requires a complete understanding of the normal process of pancreas development: an objective greatly aided by CRISPR-Cas9 genome-editing technology. First identified as the adaptive immune system of bacteria, CRISPR-Cas9 uses RNA to specifically target a DNA endonuclease to the genome, generating a double-strand break that can either be repaired by the error-prone NHEJ or via HDR. From the first demonstration that CRISPR-Cas can be programmed to cleave DNA in 2012, the field has advanced fast and now includes examples of targeting in many model organisms as well as gene knockout or reporter hPSC lines that will aid in the production of specific cell types, such as pancreatic β-cells.

Keywords CRISPR-Cas9 • Human pluripotent stem cells • Diabetes • β-cell • Endoderm • Pluripotency • Developmental biology

Abbreviations

AAV	Adeno-associated virus
Cas	CRISPR-associated
CRISPR	Clustered regularly-interspaced short palindromic repeats
dCas9	Dead Cas9
DHFR	Dihydrofolate reductase

N.A.J. Krentz (✉) • F.C. Lynn (✉)
Diabetes Research Group, Child and Family Research Institute,
950 28th Ave. W., Vancouver, BC, Canada, V5Z 4H4

Departments of Surgery/Cellular and Physiological Sciences, University of British Columbia, Vancouver, BC, Canada
e-mail: nkrentz@alumni.ubc.ca; francis.lynn@ubc.ca

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_8

DSB	Double-strand break
gRNA	Guide RNA
GUIDE-seq	Genome-wide unbiased identification of double-strand breaks enabled by sequencing
HDR	Homology-directed repair
hESC	Human embryonic stem cells
iCRISPR	Induced CRISPR
iPSC	Induced pluripotent stem cell
MODY4	Mature-onset diabetes of the young 4
Ngn3	Neurogenin3
NHEJ	Nonhomologous end-joining
PAM	Protospacer-adjacent motif
Pdx1	Pancreatic and duodenal homeobox 1
TALEN	Transcription activator-like effector nuclease
tracrRNA	Trans-activating crRNA
ZFN	Zinc-finger nuclease

Importance of Genome-Editing Technologies for Regenerative Medicine

Human embryonic stem cells (hESCs) are pluripotent cells found in the inner cell mass of preimplantation embryos. For reasons of their unlimited capacity for self-renewal and ability to form all cell types, hESCs are an excellent source of potential cells for the regenerative medicine approach for the treatment of many diseases, including diabetes. As diabetes is caused by the loss of the insulin-producing β-cells of the pancreas, simply replacing these cells represents a potential cure for diabetes. Although great progress has been made and current protocols are tantalizingly close [1, 2], the final objective of making functional β-cells has not been realized, likely because of the lack of an easy readout to quickly evaluate the efficiency of protocols, an incomplete understanding of normal mammalian pancreas development, and the subtle differences between mouse and human β-cell maturation. The generation of knock-in reporter and knockout mutant hESC lines will greatly aid our understanding of human β-cell differentiation and propel us toward the final goal of curing diabetes using stem-cell derived endocrine cells.

Previously, creating these lines in hESCs was difficult because of the low rate of homologous recombination. However, the advent of genome-editing technologies such as zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein (Cas) allows for the site-specific generation of double-strand DNA breaks to increase the frequency of homologous recombination [3–5]. By far, the most popular of these approaches is the CRISPR-Cas system because of the relative ease, specificity, and efficiency of this technology. The CRISPR-Cas

system uses RNA to direct a Cas nuclease to a genome sequence. First discovered in the adaptive immune system of bacteria, CRISPR-Cas was quickly adapted for genome editing and has fast become a "must-have" for every stem cell researcher.

Identification of CRISPR Array: The Real Heroes

In 1987, an unusual structure of repetitive DNA that was first described in *Escherichia coli* [6]. This genomic sequence consists of repeats 24 to 47 bp long with unique intervening spacer sequences 27 to 72 bp in length [7] and were named clustered regularly interspaced palindromic repeats (CRISPR) by Jansen and colleagues in 2002 [7]. The repeats are exclusively found in prokaryotes, with approximately 50 % of all bacterial and nearly all sequenced Archaea genomes containing these sequences [8]. Despite their abundance, the source and significance of these repeats remained a mystery. The first clue of its biological role was the presence of genes encoding both a DNA helicase and polymerase, suggesting that it may involve DNA repair [9].

In 2005, three independent research groups discovered that the unique spacer sequences are derived from invading bacteriophage genomes [10–12], which led to the proposal that the sequences are part of a bacterial adaptive immune system [13–15] that defends the host against invading pathogens via an RNA interference-like mechanism [16]. It is now clear that when foreign DNA is detected in prokaryotic cells the CRISPR system responds by integrating short fragments of the foreign DNA into the host chromosome at the 5′-end of the CRISPR locus [13, 16, 17], and the resulting spacers act as "vaccination cards" serving as a genetic record of prior encounters with foreign DNA [13, 17, 18] (Fig. 1).

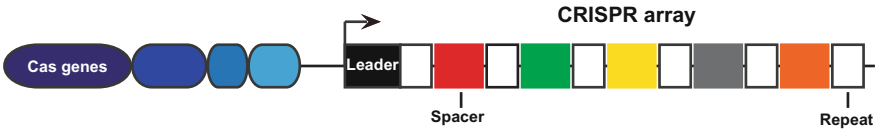

Fig. 1 Structure of the CRISPR array. Unique spacer sequences (*colored boxes*) derived from invading phage genomes are separated by repeats 24–47 bp in length. Adjacent to the spacer repeat array is an AT-rich leader sequence (*black box*) containing promoter elements and binding sites for regulatory proteins. Upon detection of foreign DNA in prokaryotic cells, short fragments of the foreign DNA are integrated as new spacer sequences at the 5′-end of the CRISPR locus

In a unique collaboration between academia and industry, the hypothesis that these repetitive sequences of DNA are part of an adaptive immune system was investigated by challenging the bacteria *Streptococcus thermophilus* with two different phages isolated from yogurt and screening the resulting cultures for bacteriophage-resistant mutants [13]. All the nine isolated mutants contained between one to four new spacer sequences added to the leader end that were derived from either the sense or antisense strand of phage DNA. It was determined that the unique spacer sequences were critical for the resistance to infection as single-nucleotide differences between the spacer and phage did not confer resistance. Inserting or removing pathogen-specific spacer sequences from the CRISPR loci resulted in improved or decreased resistance to infections [13]. Thus, the number of spacers in the CRISPR locus that are homologous to the invading DNA dictates the sensitivity of the host to a challenging phage [10]. Although successive phage challenges results in the addition of spacer sequences at the leader end of loci, there is occasional loss of repeat spacer units, usually of those found toward the trailer end of the locus [14]; however, the addition of spacers appears to occur more frequently than the loss of repeat spacer units [11]. The discovery of these repeats and their apparent role in the adaptive immune system of bacteria led to a great interest in understanding their structure and mechanism.

Structure of CRISPR Arrays

The CRISPR locus is flanked by a diverse set of *CRISPR-associated* (*cas*) genes. A systematic study aimed at characterizing proteins in the vicinity of the CRISPR locus determined that the loci are large, heterogeneous, and complex, containing up to 20 different *cas* genes [19]. The defining characteristic of the locus is the series of direct repeats that are separated by unique spacer sequences of a similar length [16, 20, 21], the number of which can vary from a few to several hundred. The spacer repeats array is generally flanked by an AT-rich sequence [7] that is thought to contain both promoter elements [22–24] and regulatory protein-binding sites [23, 24]. As the spacer sequences closest to the leader are more diverse than those in the trailer [7, 11], there is a polarity that is defined by the leader. Sequences of foreign DNA that are selected for integration are called protospacers and are flanked by sequence-specific protospacer-adjacent motifs (PAMs) [10, 14].

Because of the complexity and diversity of the CRISPR-Cas system, a classification system of three major types (types I, II, III) was proposed based on phylogeny of *cas* genes, sequence/organization of repeats, and the architecture of loci [25]. Each of the three main types of systems differs in how they mechanistically prevent infections, and more than one type can be active at a time in an organism [26]. Space constraints require this book chapter to focus solely on type II systems and their use in genome editing; however, there are several great reviews detailing the major CRISPR variants [27, 28].

The type II system has only been found in bacteria [25] and is the simplest of all types of microbial CRISPRs [28–30]. It is the most compact with only four *cas* genes: *cas1*, *cas2*, *cas9*, and either *csn2* (type II-A) or *cas4* (type II-B). Cas1 and Cas2 are the only proteins that are universally conserved across all types of CRISPR-Cas systems [16, 19]. Two defining features of the type II system is the large multi-functional protein Cas9 and the requirement of the trans-acting CRISPR-associated RNA (tracrRNA) that is encoded upstream and on the opposite strand of the CRISPR-Cas locus [31].

Mechanism of Adaptive Immunity

There are three stages of CRISPR-mediated adaptive immunity: spacer acquisition, crRNA maturation, and interference [30, 32, 33] (Fig. 2).

Spacer acquisition occurs when a short protospacer sequence of foreign DNA is incorporated into the CRISPR array as a new spacer [18], serving as a genetic record of previous infections [13]; this requires the protospacer-adjacent motif (PAM) sequence, a four- or five-nucleotide sequence that is directly 5′ to the protospacer on the target DNA strand and not included in the spacer itself [10, 14, 15, 17] (Fig. 2a). PAM sequences were first discovered as a conserved sequence adjacent to protospacers in phage DNA using in silico analysis [10], and the PAM sequence depends on the CRISPR-Cas variant [34]. The PAM is important for the discrimination of self versus non-self [35] by distinguishing potential target sequences from CRISPR spacers in the loci [36, 37] as the spacers do not contain the required adjacent PAM sequence [10, 14]. In addition, the PAM sequence is necessary for efficient target binding [38], for spacer acquisition as integration machinery recognizes the PAM [39], and for target interference [14, 17, 38].

The strict requirement for PAM sequences in type II CRISPR-Cas systems prevents DNA cleavage from occurring when the PAM sequence is mutated [14, 38, 40–43], resulting in phages adapting to avoid CRISPR immunity when mutations occur in the PAM sequence [14]. This protection from CRISPR-mediated interference with a single mutation in the PAM sequence occurs even if the spacer and protospacer sequences have 100 % complementarity [14, 17, 38, 44].

After spacer acquisition and to confer immunity against a pathogen, the spacer sequences need to undergo maturation to produce crRNA (Fig. 2b). This process requires two distinct steps: first, the CRISPR loci is transcribed as a long precursor crRNA (pre-crRNA) from promoter sequences in leader, and second, the pre-crRNA is processed into mature crRNA by the subtype-specific enzymes Cas9 and RNAse III. The mature crRNA then hybridizes with the tracrRNA to form a dsRNA that is recognized and cleaved by RNAse III during crRNA maturation [31]. As deletion of Cas9 prevents crRNA biogenesis, Cas9 is required for crRNA maturation but its exact role in this process is unknown [31]. It is also during the crRNA maturation phase of adaptive immunity that the *cas* genes are transcribed and translated.

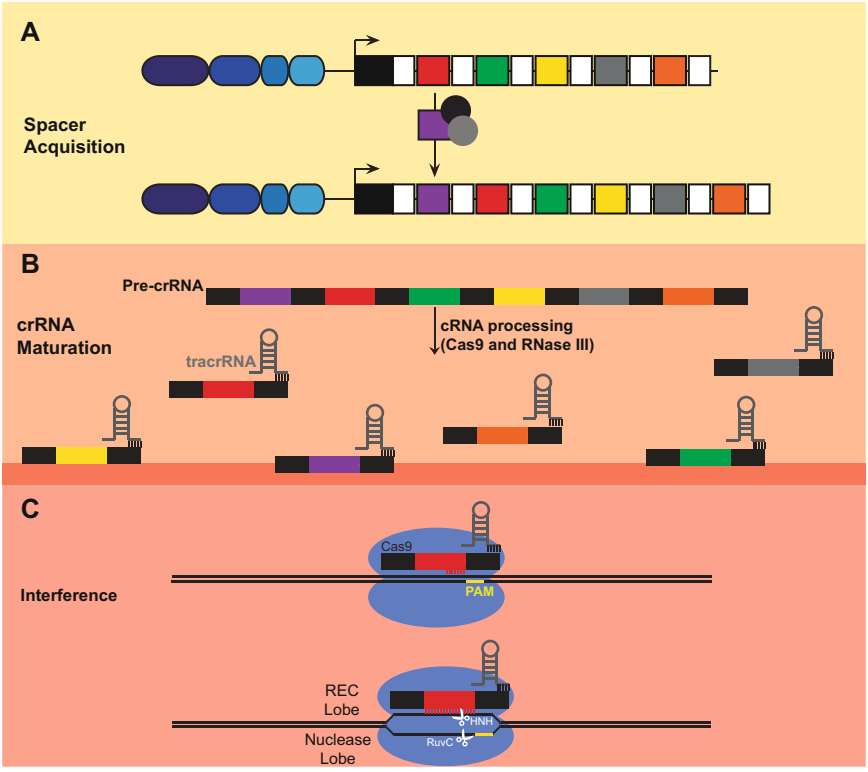

Fig. 2 The three stages of CRISPR-mediated adaptive immunity. (**a**) Spacer acquisition results in a short protospacer sequence of foreign DNA (*purple square*) being incorporated at the leader end (*black square*) of the CRISPR locus, separated from the adjacent spacer sequence (*red square*) by repeat sequences (*white square*). Cas genes are located upstream of the repeat spacer units (*blue ovals*). (**b**) crRNA maturation begins with transcription of the long precursor crRNA (pre-crRNA). The pre-crRNA is then processed into mature crRNA by Cas9 and RNase III and hybridizes with the tracrRNA to form a dsRNA. (**c**) Target sequences are cleaved during interference when a mature crRNA recognizes its complementary sequence in a process that involves an adjacent PAM sequence (*yellow*). The Cas9 enzyme has two lobes: the target recognition (REC) lobe and the nuclease lobe. Within the nuclease lobe the RuvC domain cleaves the nontarget strand and the HNH domain cleaves the target strand DNA to generate a blunt-end cleavage product

Finally, during interference, the mature crRNA "guides" cas protein(s) to cleave the complementary nucleic acid [17, 32, 33, 40, 45, 46] (Fig. 2c). In type II systems, this involves a single protein, Cas9, along with the dual-RNA heteroduplex of tracrRNA and crRNA [31, 40]. Cas9 requires both the mature crRNA and tracrRNA for DNA cleavage to occur during target interference [40] and each crRNA:tracrRNA complex can only recognize one target that is dictated by the sequence of the spacer that it contains [31, 36, 45, 47–54].

The success of the interference phase of CRISPR adaptive immunity depends on several factors including the expression of Cas9, which is important not only for

crRNA biogenesis but also for target interference [17, 31, 40]. The Cas9 protein has a bilobed architecture that consists of the target recognition and the nuclease lobes with the positively charged groove at the interface [55]. To investigate the bilobed nature of Cas9 protein, a split-Cas9 enzyme was created in which the nuclease lobe and the recognition lobe were expressed as separate peptides [56]. The two lobes do not interact on their own, but the crRNA:tracrRNA is necessary and sufficient to dimerize the nuclease and recognition lobes into an active complex [56]. Within the nuclease lobe there are two domains: the HNH domain that is responsible for DNA cleavage of the target strand and the RuvC-like domain which cleaves DNA on the nontarget strand [40, 57] (Fig. 2c). DNA cleavage occurs in the protospacer sequence to generate a blunt-end cleavage product [17].

Although the PAM sequence is required for spacer acquisition, it is also strictly required for interference as DNA cleavage can only occur when there is a PAM sequence on the target DNA, allowing for Watson–Crick base-pairing between the crRNA:tracrRNA sequence and its complementary target sequence [40]. Although Cas9 can transiently and weakly react with DNA in the absence of PAM, evidence from single-molecule-based assays suggests that Cas9 preferentially interacts and binds longer to DNA with PAM sequences [58, 59]. As the Cas9 protein is the limiting factor for Cas9-mediated DNA cleavage [60] and the DNA sequence is interrogated by Cas9-RNA beginning at the PAM sequence and proceeding directionally toward the distal end of target sequence [58], the PAM sequence is integral for the CRISPR-mediated adaptive immunity.

Adaptation of Type II CRISPR-Cas for Genome Editing

The discovery of an RNA-based DNA cleavage system in bacteria was quickly adapted for genome editing: with the type II CRISPR-Cas system from *Streptococcus pyogenes* being the most widely applied. The most common system uses a guide RNA (gRNA) that is a single RNA chimera, produced by fusion of the crRNA and tracrRNA with an addition of 20 nt to the 5′-end (protospacer region of crRNA) that is sufficient to guide Cas9 to an intended target site [40]. The gRNA was also extended at the 3′-end to improve DNA-targeting activity [40]. Thus, the gRNA in this system is about 100 nt long with a 17- to 20-nt sequence at the 5′-end that specifies DNA target sequence. Similarly to adaptive immunity, the target site is recognized by Watson–Crick base-pairing between the gRNA and target DNA strand, is mediated by Cas9 [16, 40] and requires a PAM sequence for the generation of a double-stranded break (DSB).

The DSB that is generated can be repaired in one of two ways: nonhomologous end-joining (NHEJ) or homology-directed repair (HDR). DSBs that are repaired by NHEJ can induce missense or nonsense mutations (indels) when the break occurs in an open reading frame. For HDR, a DNA template with homology arms is required and can be as short as 10–30 nt [61, 62], allowing for the use of single-stranded DNA oligonucleotides [61, 63, 64].

When using genome-editing technology, it is important to consider targeting efficiency, which is largely dependent on the cell type that is being targeted and the genomic location of the targeting site. Several studies targeting the same loci using both TALEN and CRISPR-Cas9 systems have determined that CRISPR technology is more efficient [65, 66]. Although DNA methylation has been reported to decrease efficiency when using TALENs [67], limited reports suggest that CRISPR-Cas9 may not be as sensitive to DNA methylation [68]. To improve the efficiency of HDR, nucleofection of preassembled Cas9/guide RNA complexes can be used [69]. In addition, generating KO via NHEJ can be improved using a modified gRNA structure with an extended duplex length and a mutated fourth thymine to either cytosine or guanine [70].

Reducing the Risk of Off-Target Mutagenesis When Using CRISPR-Cas9

An important consideration when undertaking genome editing is the specificity of targeting as off-target binding of Cas9 enzymes in unintended regions of the genome could lead to double-stranded breaks and induce mutations. As CRISPR-Cas9 only requires the binding of a single RNA to generate the double-stranded break, there is an increased chance of erroneous genome binding. The specificity of CRISPR-Cas9 and TALENs was compared by targeting GFP into the AAVS1 locus followed by whole-genome sequencing, determining that CRISPR-Cas9 was not only as efficient but also as specific as TALENs [71]; this was corroborated by other studies that determined that off-target mutations by Cas9 are very rare [72, 73]. The rates of mutagenesis were further studied using the technique GUIDE-seq (genome-wide, unbiased identification of DSBs enabled by sequencing), which involves sequencing of double-stranded oligodeoxynucleotides that are incorporated into DSBs generated by Cas9. Using this approach the authors investigated 13 different gRNAs in human cells and determined that off-target binding varies and is largely dependent on gRNA [74], a finding that is consistent with other reports [75, 76].

ChIP-seq data using 12 gRNAs and a catalytically dead mutant Cas9 enzyme (dCas9) revealed that off-target binding varies from as little as ten instances to greater than a thousand and is dependent on both the gRNA and the genomic regions 3′ of the PAM [77]. To determine if this binding was sufficient to cause cleavage, the authors performed whole-genome sequencing using the catalytically active form of Cas9 and found that although some off-target sites had indels at rates higher than background, mutation frequencies were significantly reduced compared to those at intended target sites [77]. The finding that off-target binding may not result in cleavage led to the proposal of a two-state model, where seed match triggers binding but extensive pairing is required for cleavage [78]. Consistently, single-particle tracking of Cas9 determined that off-target binding in living cells is short lived [79] and that Cas9 cleaves off-target DNA in a sequence-dependent manner that is sensitive to the number, position, and distribution of the mismatches between gRNA and genomic sequence [80].

Because much of the off-target activity of Cas9 appears to be gRNA dependent, care must be taken to select unique target sequences in the genome [81], and there

are several Web-based tools to aid in the selection of potential targeting sequences [80, 82–85]. Early models suggested that the first seven base pairs of the gRNA seed sequence were important for binding, but high-throughput sequencing of eight gRNAs determined that specificity extends into the 7- to 12-bp region as well [86], suggesting that particular attention must be paid to the first 12 bp proximal to the PAM when designing gRNAs [87]. Using truncated gRNAs of fewer than 20 nucleotides decreases off-target mutagenesis by 5000-fold [74, 88] but likely reduces activity [86]. In addition, the concentration of gRNA can be titrated to minimize off-target effects [80], as high concentrations of gRNA can lead to off-target cleavage even with mismatches that are near or within the PAM [86, 88].

Another consideration for target design is the presence of the PAM sequence, as its abundance in the genome adjacent to the complementary gRNA sequence will increase the likelihood of off-target cleavage [58]. To address this, a side-by-side comparison of the *Streptococcus pyogenes* Cas9 that used the PAM sequence "NGG" was compared to *Streptococcus thermophilus*, which requires a longer PAM. Although these two enzymes were comparable for target cleavage, the Cas9 that required a longer PAM sequence had significantly reduced off-target mutagenesis [89]. Further characterization of other Cas9 enzymes or improved design of gRNA in the future will likely reduce the risk of off-target Cas9 recognition.

To improve specificity, a nickase that only cuts a single DNA strand can be used [40], such as the D10A Cas9 enzyme that has a mutation in the RuvC domains which does not impair binding to DNA but results in a single-strand endonuclease that is highly specific [40, 81, 90–92]. Using two nickases reduces off-target effects by 50- to 1500-fold in cell lines [81, 90, 91]. Unfortunately, this is at the cost of efficiency as it can be difficult to find two PAM sequences in close proximity. A similar approach is to use RNA to guide two Fok1 nucleases that must dimerize before cutting can occur [93, 94]. This approach reduces the off-target mutagenesis to an undetectable level when using deep-sequencing analysis and results in 140-fold higher specificity in human cells [94]. It has also been used to delete large genomic regions of 10 kb to 15 mega bases [91, 95, 96] as well as to generate mutant mice by microinjection into mouse zygotes [97].

Alternate Uses of the CRISPR-Cas9 System

The CRISPR-Cas9 system can also be used to activate or repress transcription in an inducible manner. Expression of endogenous human genes can be increased by using a dCas9 fused to VP64 (4xVP16) transcriptional activation domain [98]. Similarly, the CRISPR-On system uses RNA to guide dCas9 fused to 3xVP16 (dCas9-VP48), a system that was sufficient to activate reporter genes [99]. Using epigenetic modifiers, such as the histone acetyltransferase p300, can also enhance dCas9-VP64 induction of transcription [100]. Conversely, transcription can also be repressed by using a dCas9 fused to the KRAB-repressive domain, preventing RNA polymerase binding to promoter sequences or acting as a transcription terminator by blocking RNA polymerase [101]. Further studies suggest that dCas9-KRAB induces H3K9me3 at enhancer regions,

resulting in decreased accessibility of genomic sequences [102]. This system is sensitive to even a few mismatches [103], making it highly specific to repress transcription of only that gene of interest [103, 104]. Similarly, the expression of genes can be turned off using CRISPRi, which also uses dCas9 to block transcription via interference with transcriptional elongation, RNA polymerase binding, or transcription factor binding [105]. These strategies have also been amended for genome-scale induction or repression of genes and libraries of up to 10 gRNAs per TSS have been generated [103].

To allow for conditional control of CRISPR-Cas system, a light-activated form of Cas9 was generated by inserting a caged lysine amino acid that renders Cas9 inactive. Upon exposure to light, the caging is reversed and Cas9 becomes active [106]. Another inducible genome-editing technique is called iCRISPR, involving doxycycline-inducible expression of Cas9 in hESCs. With transfection of individual gRNAs, double- or triple-gene knockout can be generated with 10 % efficiency or knock-in clones can be created by providing a ssDNA template [107]. An hESC line with Flpe-ER^{T2} knocked in to the AAVS1 locus allows for easy inducible knockout: the exon of interest is flanked with Flp/FRT sites using CRISPR-Cas9, and upon addition of 4-OHT, flippase enters the nucleus and removes the frt flanked exon [108].

In an attempt to discover other Cas proteins that will allow for simultaneous targeting at one time, the Cas9 protein from *Neisseria meningitides* was found to target human and bacterial cells with high efficiency [90]. In search of a smaller Cas9 protein to package within a single adeno-associated virus (AAV) vector, six other Cas9 orthologues were investigated and the Cas9 protein from *Staphylococcus aureus* was found to be as efficient although being 1 kb smaller [109]. Using this strategy, the authors targeted the gene Pcsk9 and achieved indels in greater than 40 % of liver cells within 1 week of a single injection of AAV [109]. Finally, another study investigated a different family of proteins containing 16 members called Cpf1. Two of these proteins from *Acidaminococcus* and *Lachnospiraceae* were found to edit the genome of human cells with high efficiency; however, they differ from other Cas9 proteins as they do not need a tracrRNA, have a T-rich PAM, and cleave DNA via a staggered double-stranded break [110], generating sticky ends that could improve the efficiency of knocking in DNA. The addition of these other Cas proteins will allow for multiplex targeting and in vivo targeting, including targeting brain [111, 112], liver [112, 113], and immune cells [112].

Using CRISPR-Cas9 Approaches in Model Organisms

Our understanding of basic cell biology and embryonic development has been greatly aided by the study of model organisms and the adaptation of genome editing to these organisms will serve to identify new genes in developmental processes. Within the plant kingdom, CRISPR-Cas has been used to modify the genomes of *Arabidopsis*, tobacco, sorghum, rice, and sweet orange [114]. In *Saccharomyces cerevisiae*, constitutive Cas9 expression paired with transient gRNA expression increased the rates of homologous recombination to nearly

100 % [115]. Two transgenic strains of *Drosophila* have been generated wherein Cas9 is expressed ubiquitously using U6 or in the germline using *nanos* promoter, allowing for germline transition of mutations in about 60 % of progeny [75]. Similarly, *Caenorhabditis elegans* has also been targeted [116, 117]. CRISPR-Cas9 has been used in zebrafish to generate chromosomal deletions or inversions [118], biallelic modifications [119], and to target multiple genes [120] with similar efficiencies as other genome-editing strategies [121]. In mammals, CRISPR-Cas9 has been used to simultaneously target up to five genes with high efficiency in mouse embryos by co-injecting Cas9 mRNA and gRNA [122, 123]. Similarly, cynomolgus monkeys have been targeted by co-injection of Cas9 mRNA and gRNA at the one-cell stage of embryogenesis [124]. The multifaceted nature of the CRISPR-Cas9 system combined with its adoption in multiple model organisms highlights the importance and utility of genome editing in many research programs.

Generating β-Cells from Human Pluripotent Stem Cells for Diabetes Treatment

Regenerative medicine approaches to produce transplantable cells for disease treatment have shown much promise in many diseases including diabetes. Diabetes results from the loss or dysfunction of pancreatic β-cells located in the islets of Langerhans. In 2000, Shapiro and colleagues described a glucocorticoid-free immunosuppressive therapy for islet transplantation that provided insulin independence in patients with type 1 diabetes [125]. The benefits of islet transplantation at reducing secondary complications and the shortage of human donor islets have encouraged research focused on the generation of large numbers of transplantable β-cells. One potential approach is the differentiation of human embryonic stem cells (hESC). Current differentiation protocols result in endocrine progenitor cells that fully differentiate and mature into insulin-secreting β-cells upon transplantation into immunodeficient mice [126–129]. We are able to produce immature endocrine cells in culture [1, 2], but currently no differentiation protocol has been successful at producing fully functioning glucose-sensing insulin secreting β-cells in vitro. To produce mature cells for transplantation in vitro, we will require a detailed understanding of the developmental network that regulates the formation of pancreatic β-cells.

Genome-editing technologies such as CRISPR-Cas9 provide a great opportunity to improve our understanding of this process. Traditionally, targeting pluripotent cells was difficult [5] because it required homologous recombination, a rare event in human pluripotent stem cells [130]. However, the use of CRISPR-Cas9 improved the efficiency of targeting in induced pluripotent stem cells (iPSC) to 2–4 % [131], and this was further increased to 51–79 % by using fluorescent markers to select Cas9-expressing cells [65], making generating knockout or knock-in reporter lines more feasible in stem cell research.

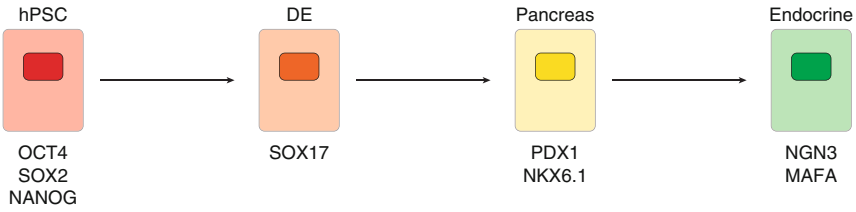

Fig. 3 Schematic of CRISPR-Cas9-targeted genes in the developmental pathway from stem cell to beta cell. Reporter hPSC lines have been generated for the pluripotency genes OCT4, SOX2, and NANOG, the definitive endoderm (DE) gene SOX17, and the pancreas gene PDX1. hPSC lines that allow for the induction of the definitive endoderm gene SOX17, pancreatic genes PDX1 and NKX6.1, and the endocrine gene MAFA have also been generated; CRISPR-Cas9 was used to repress expression of the pluripotency gene OCT4. Knockout mutant lines have been generated for the endocrine genes NGN3

The first step in the differentiation protocol toward any somatic cell type is to lose expression of pluripotency genes such as OCT4, SOX2, and NANOG (Fig. 3). OCT4/POU5F1 is an important member of the pluripotency network as loss of Oct4 in mouse embryos results in a loss of pluripotent cells [132], and Oct4 is required for the formation of all embryonic germ layers in vitro and in vivo [133]. Owing to its important role in pluripotency and differentiation, many reporter lines have been generated to understand the role of OCT4 in early fate decisions in human pluripotent cells. While this was traditionally done using transgenic approaches [134–136], recent advances in genome editing have allowed for the knock-in of fluorescent proteins directly into the OCT4 locus of hPSCs [66, 137, 138]. Another important member of the pluripotency network is Sox2, which works together with Oct4 to regulate the transcription of many genes [139]. Using homologous recombination mediated by an adeno-associated virus, the SOX2 coding region was replaced with eGFP-SV40-NeopA in H9 hESCs, which did not affect the formation of endoderm [140]. Another study used homologous recombination to insert eGFP into exon 1 of NANOG in HUES-1 and HUES-3 hESCs and determined that eGFP expression overlapped with NANOG immunofluorescence and was lost upon differentiation [141].

The important role of OCT4 both in maintaining pluripotency and in directing differentiation resulted in the generation of strategies to control transcription of OCT4 during hPSC differentiation to the endoderm lineage, the germ layer from which many tissues, including the pancreas, are derived [142]. Using a CRISPRi system wherein dCas9 is fused to the transcriptional repressor KRAB, a gRNA was used to repress transcription of the OCT4a gene [143]. Conversely, fusing dCas9 to the activation domain VP16 was used to increase OCT4 expression up to 70-fold [144]. A similar strategy with dCas9-VP64 was sufficient to induce the expression of SOX17 [143], a protein important for endoderm formation. This technique paired with a H9 reporter line that marks SOX17-expressing cells with GFP [145] will greatly improve our understanding of early fate decisions and how SOX17 regulates this process.

A subset of endoderm cells activates the transcription factor pancreatic and duodenal homeobox 1 (PDX1), giving rise to the pancreas and gut. Mice null for *Pdx1*

have pancreatic agenesis [146] that is thought to be the result of a reduction in progenitor cell proliferation [147]. As with mouse pancreas development, PDX1 is required for human pancreas formation [148], and patients with a single nucleotide deletion in *PDX1* also have pancreatic agenesis [149]. Accordingly, mutations in *PDX1* have been linked to maturity-onset diabetes of the young (MODY4) [150] and permanent neonatal diabetes [151].

To investigate the role of PDX1 in human pancreas development, the doxycycline-inducible Cas9 (iCRISPR) line, where gRNA transfection in combination with doxycycline treatment improves biallelic knockout efficiency to 17–67 % [107], was used to generate PDX1-eGFP reporter lines [152], an important resource for understanding specification of the pancreatic lineage from definitive endoderm cells. Another important tool is to activate expression of pancreas-specific genes in an inducible manner. PDX1 expression was controlled using gRNAs and dCas9 fused to a dihydrofolate reductase (DHFR)-derived destabilization domain that is only stabilized upon addition of Trimethoprim [144]. This approach was sufficient to simultaneously activate PDX1 and NKX6.1 expression during hPSC differentiation in combination with doxycycline induction of MAFA [144], highlighting the feasibility of controlling the activation or repression of transcriptional programs to improve the efficiencies of current differentiation protocols.

The majority of endocrine differentiation occurs in mice between E13.5 and E15.5 in a process known as the "secondary transition." Neurogenin3 (Ngn3), a basic-helix-loop-helix transcription factor that is markedly induced toward the end of the secondary transition, marks the cells fated to become endocrine cells [153] and is required for the formation of all endocrine cell types [154]. Moreover, ectopic expression of Ngn3 is sufficient to generate all islet cell types in vivo [153, 155, 156]. A subset of cells within the pancreatic epithelium will activate Ngn3 expression, migrate away from the epithelium, and begin to activate the lineage-specific gene expression cascade (reviewed by Pan and Wright [157]). However, the process that governs which progenitor cells activate Ngn3 and why Ngn3 is activated is not well understood. In human development, by gestational week 8 NGN3 protein expression is detected within a subset of pancreatic epithelial cells [158]. Neurog3 is required for the formation of all endocrine cell types in the mouse pancreas [153] but patients with biallelic mutations in NEUROG3 are born with circulating C-peptide levels [159], suggesting the NEUROG3 may not be as important for the formation of human endocrine pancreas. To investigate the role of NEUROG3 in human pancreas formation, CRISPR-Cas9 was used to generate hPSCs with heterozygous or homozygous mutations in NEUROG3, and the formation of endocrine cells was determined using directed differentiation protocols [160]. From this study it was determined the NEUROG3 is required for endocrine formation during human pancreas development. With the relative ease of generating biallelic mutations in hPSCs using CRISPR-Cas9, it will be possible to determine the function of many new genes in human endocrine cell differentiation as well as to understand some of the mutations that cause monogenic diabetes and to learn how genetic variants drive the development of type 1 and 2 diabetes [161].

Conclusion

Regenerative medicine approaches to treat diseases such as diabetes by creating an unlimited source of transplantable cells from hESCs or iPSCs has been of great interest to many researchers. Manipulating the genome of pluripotent cells was previously difficult because of the low rates of homologous recombination, but new genome-editing technologies such as CRISPR-Cas now allow the rapid creation of knockout mutant lines and knock-in reporter lines. First discovered in the adaptive immune system of bacteria, CRISPR-Cas has fast become essential for every stem cell researcher. Although the use of genome editing in investigating pancreas development, monogenic forms of diabetes, or the metabolic disorder itself is still in its infancy, the relative ease and accessibility of CRISPR-Cas9 promises to lead to many new advances in these areas of research.

References

1. Rezania A, Bruin JE, Arora P, Rubin A, Batushansky I, Asadi A, et al. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. Nat Biotechnol. 2014;32(11):1121–33.
2. Pagliuca FW, Millman JR, Gürtler M, Segel M, Van Dervort A, Ryu JH, et al. Generation of functional human pancreatic β cells in vitro. Cell. 2014;159(2):428–39.
3. Ding Q, Lee Y-K, Schaefer EAK, Peters DT, Veres A, Kim K, et al. A TALEN genome-editing system for generating human stem cell-based disease models. Cell Stem Cell. 2012;12(2):238–51.
4. Gaj T, Gersbach CA, Barbas III CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 2013;31(7):397–405.
5. Hockemeyer D, Jaenisch R. Gene targeting in human pluripotent cells. Cold Spring Harb Symp Quant Biol. 2010;75:201–9.
6. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. J Bacteriol. 1987;169(12):5429–33.
7. Jansen R, Embden JDAV, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol. 2002;43(6):1565–75.
8. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science. 2010;327(5962):167–70.
9. Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. Nucleic Acids Res. 2002;30(2):482–96.
10. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 2005;151(Pt 8):2551–61.
11. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology. 2005;151(Pt 3):653–63.
12. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol. 2005;60(2):174–82.
13. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315(5819):1709–12.

14. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J Bacteriol. 2008;190(4):1390–400.
15. Horvath P, Romero DA, Coute-Monvoisin AC, Richards M, Deveau H, Moineau S, et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. J Bacteriol. 2008;190(4):1401–12.
16. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol Direct. 2006;1:7.
17. Garneau JE, Dupuis M-È, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468(7320):67–71.
18. Heler R, Marraffini LA, Bikard D. Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. Mol Microbiol. 2014;93(1):1–9.
19. Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput Biol. 2005;1(6):e60.
20. Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. Genome Biol. 2007;8(4):R61.
21. Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics. 2007;8(1):172.
22. Hale CR, Majumdar S, Elmore J, Pfister N, Compton M, Olson S, et al. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. Mol Cell. 2012;45(3):292–302.
23. Pougach K, Semenova E, Bogdanova E, Datsenko KA, Djordjevic M, Wanner BL, et al. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. Mol Microbiol. 2010;77(6):1367–79.
24. Pul Ü, Wurm R, Arslan Z, Geißen R, Hofmann N, Wagner R. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. Mol Microbiol. 2010;75(6):1495–512.
25. Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR–Cas systems. Nat Rev Microbiol. 2011;9(6):467–77.
26. Carte J, Christopher RT, Smith JT, Olson S, Barrangou R, Moineau S, et al. The three major types of CRISPR-Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. Mol Microbiol. 2014;93(1):98–112.
27. Hochstrasser ML, Doudna JA. Cutting it close: CRISPR-associated endoribonuclease structure and function. Trends Biochem Sci. 2015;40(1):58–66.
28. Barrangou R, Marraffini LA. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. Mol Cell. 2014;54(2):234–44.
29. Gasiunas G, Sinkunas T, Siksnys V. Molecular mechanisms of CRISPR-mediated microbial immunity. Cell Mol Life Sci. 2013;71(3):449–65.
30. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012;482(7385):331–8.
31. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011;471(7340):602–7.
32. van der Oost J, Westra ER, Jackson RN, Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR–Cas systems. Nat Rev Microbiol. 2014;12(7):479–92.
33. Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science. 2008;322:1843–5.
34. Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009;155(3):733–40.

35. Marraffini LA, Sontheimer EJ. Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature. 2010;463(7280):568–71.
36. Jore MM, Lundgren M, van Duijn E, Bultema JB, Westra ER, Waghmare SP, et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nat Struct Mol Biol. 2011;18(5):529–36.
37. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends Biochem Sci. 2009;34(8):401–7.
38. Semenova E, Jore MM, Datsenko KA. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc Natl Acad Sci U S A. 2011;108(25):10098–103.
39. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucleic Acids Res. 2012;40(12):5569–76.
40. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816–21.
41. Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res. 2013;41(20):e188.
42. Westra ER, Semenova E, Datsenko KA, Jackson RN, Wiedenheft B, Severinov K, et al. Type I-E CRISPR-Cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. PLoS Genet. 2013;9(9):e1003742.
43. Fineran PC, Gerritzen MJH, Suarez-Diez M, Kunne T, Boekhorst J, van Hijum SAFT, et al. Degenerate target sites mediate rapid primed CRISPR adaptation. Proc Natl Acad Sci U S A. 2014;111(16):E1629–38.
44. Gudbergsdottir S, Deng L, Chen Z, Jensen JVK, Jensen LR, She Q, et al. Dynamic properties of the *Sulfolobus* CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. Mol Microbiol. 2010;79(1):35–49.
45. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008;321(5891):960–4.
46. Hale C, Kleppe K, Terns RM, Terns MP. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. RNA. 2008;14(12):2572–9.
47. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev. 2008;22(24):3489–96.
48. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. Science. 2010;329(5997):1355–8.
49. Sashital DG, Jinek M, Doudna JA. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. Nat Struct Mol Biol. 2011;18(6):680–7.
50. Gesner EM, Schellenberg MJ, Garside EL, George MM, MacMillan AM. Recognition and maturation of effector RNAs in a CRISPR interference pathway. Nat Struct Mol Biol. 2011;18(6):688–92.
51. Lintner NG, Kerou M, Brumfield SK, Graham S, Liu H, Naismith JH, et al. Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). J Biol Chem. 2011;286(24):21643–56.
52. Hatoum-Aslan A, Maniv I, Marraffini LA. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. Proc Natl Acad Sci U S A. 2011;108(52):21218–22.
53. Richter H, Zoephel J, Schermuly J, Maticzka D, Backofen R, Randau L. Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis*. Nucleic Acids Res. 2012;40(19):9887–96.
54. Nam KH, Haitjema C, Liu X, Ding F, Wang H, DeLisa MP, et al. Cas5d protein processes Pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. Structure. 2012;20(9):1574–84.
55. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014;156(5):935–49.

56. Wright AV, Sternberg SH, Taylor DW, Staahl BT, Bardales JA, Kornfeld JE, et al. Rational design of a split-Cas9 enzyme complex. Proc Natl Acad Sci U S A. 2015;112(10):2984–9.
57. Gasiunas G, Barrangou R. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci U S A. 2012;109(39):E2579–86.
58. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2015;507(7490):62–7.
59. Szczelkun MD, Tikhomirova MS, Sinkunas T, Gasiunas G, Karvelis T, Pschera P, et al. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc Natl Acad Sci U S A. 2014;111(27):9798–803.
60. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. eLife. 2013;2, e00471.
61. Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug II RG, et al. In vivo genome editing using a high-efficiency TALEN system. Nature. 2013;490(7422):114–8.
62. Campbell JM, Hartjes KA, Nelson TJ, Xu X, Ekker SC. New and TALENted genome engineering toolbox. Circ Res. 2013;113(5):571–87.
63. Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res. 2010;38(15):e152.
64. Wefers B, Meyer M, Ortiz O. Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. Proc Natl Acad Sci U S A. 2013;110(10):3782–7.
65. Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. Cell Stem Cell. 2013;12(4):393–4.
66. Krentz NAJ, Nian C, Lynn FC. TALEN/CRISPR-mediated eGFP knock-in add-on at the OCT4 locus does not impact differentiation of human embryonic stem cells towards endoderm. PLoS One. 2014;9(12):e114275.
67. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods. 2010;8(1):74–9.
68. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281–308.
69. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife. 2014;3, e04766.
70. Dang Y, Jia G, Choi J, Ma H, Anaya E, Ye C, et al. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. Genome Biol. 2015;16:280.
71. Smith C, Gore A, Yan W, Abalde-Atristain L, Li Z, He C, et al. Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. Cell Stem Cell. 2014;15(1):12–3.
72. Veres A, Gosis BS, Ding Q, Collins R, Ragavendran A, Brand H, et al. Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell. 2014;15(1):27–30.
73. Suzuki K, Yu C, Qu J, Li M, Yao X, Yuan T, et al. Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. Cell Stem Cell. 2014;15(1):31–6.
74. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol. 2014;33(2):187–97.
75. Kondo S, Ueda R. Highly improved gene targeting by germline-specific Cas9 expression in *Drosophila*. Genetics. 2013;195(3):715–21.
76. Sebo ZL, Lee HB, Peng Y, Guo Y. A simplified and efficient germline-specific CRISPR/Cas9 system for *Drosophila* genomic engineering. Fly. 2014;8(1):52–7.
77. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat Biotechnol. 2014;32(7):677–83.
78. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB, et al. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol. 2014;32(7):670–6.

79. Knight SC, Xie L, Deng W, Guglielmi B, Witkowsky LB, Bosanac L, et al. Dynamics of CRISPR-Cas9 genome interrogation in living cells. Science. 2015;350(6262):823–6.
80. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31(9):827–32.
81. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. 2014;24(1):132–41.
82. Park J, Bae S, Kim J-S. Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites. Bioinformatics. 2015;31(24):4014–6.
83. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. 2014;15(12):554.
84. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol. 2014;32(12):1262–7.
85. Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, et al. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. Genetics. 2014;196(4):961–71.
86. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol. 2013;31(9):839–43.
87. Cradick TJ, Fine EJ, Antico CJ, Bao G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res. 2013;41(20):9584–92.
88. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol. 2014;32(3):279–84.
89. Muller M, Lee CM, Gasiunas G, Davis TH, Cradick TJ, Siksnys V, et al. *Streptococcus thermophilus* CRISPR-Cas9 systems enable specific editing of the human genome. Mol Ther. 2016;24:636–44.
90. Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat Methods. 2013;10(11):1116–21.
91. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013;154(6):1380–9.
92. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, et al. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Methods. 2014;11(4):399–402.
93. Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol. 2014;32(6):569–76.
94. Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat Biotechnol. 2014;32(6):577–82.
95. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819–23.
96. Fujii W, Kawasaki K, Sugiura K, Naito K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. Nucleic Acids Res. 2013;41(20):e187.
97. Hara S, Tamano M, Yamashita S, Kato T, Saito T, Sakuma T, et al. Generation of mutant mice via the CRISPR/Cas9 system using FokI-dCas9. Sci Rep. 2015;5:11221.
98. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. Nat Methods. 2013;10(10):977–81.
99. Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res. 2013;23(10):1163–71.
100. Nissim L, Perli SD, Fridkin A, Perez-Pinera P, Lu TK. Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. Mol Cell. 2014;54(4):698–710.

101. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. 2013;41(15):7429–37.
102. Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat Methods. 2015;12(12):1143–9.
103. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, et al. Genome-scale CRISPR-mediated control of gene repression and activation. Cell. 2014;159(3):647–61.
104. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013;154(2):442–51.
105. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152(5):1173–83.
106. Hemphill J, Borchardt EK, Brown K, Asokan A, Deiters A. Optical control of CRISPR/Cas9 gene editing. J Am Chem Soc. 2015;137(17):5642–5.
107. González F, Zhu Z, Shi Z-D, Lelli K, Verma N, Li QV, et al. An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell. 2014;15(2):215–26.
108. Chen Y, Cao J, Xiong M, Petersen AJ, Dong Y, Tao Y, et al. Engineering human stem cell lines with inducible gene knockout using CRISPR/Cas9. Cell Stem Cell. 2015;17(2):233–44.
109. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. Nature. 2015;520(7546):186–91.
110. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015;163(3):759–71.
111. Swiech L, Heidenreich M, Banerjee A, Habib N, Li Y, Trombetta J, et al. In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. Nat Biotechnol. 2015;33(1):102–6.
112. Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell. 2014;159(2):440–55.
113. Xue W, Chen S, Yin H, Tammela T, Papagiannakopoulos T, Joshi NS, et al. CRISPR-mediated direct mutation of cancer genes in the mouse liver. Nature. 2014;514(7522):380–4.
114. Jia H, Wang N. Targeted genome editing of sweet orange using Cas9/sgRNA. PLoS One. 2014;9(4):e93806.
115. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 2013;41(7):4336–43.
116. Lo TW, Pickle CS, Lin S, Ralston EJ, Gurling M. Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics. 2013;195(2):331–48.
117. Zhao P, Zhang Z, Ke H, Yue Y, Xue D. Oligonucleotide-based targeted gene editing in *C. elegans* via the CRISPR/Cas9 system. Cell Res. 2014;24(2):247–50.
118. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, et al. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res. 2013;41(14):e141.
119. Blitz IL, Biesinger J, Xie X, Cho KWY. Biallelic genome modification in F_0 *Xenopus tropicalis* embryos using the CRISPR/Cas system. Genesis. 2013;51(12):827–34.
120. Guo X, Zhang T, Hu Z, Zhang Y, Shi Z, Wang Q, et al. Efficient RNA/Cas9-mediated genome editing in *Xenopus tropicalis*. J Embryol Exp Morphol. 2014;141(3):707–14.
121. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. 2013;31(3):227–9.
122. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell. 2013;153(4):910–8.

123. Yang H, Wang H, Jaenisch R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. Nat Protoc. 2014;9(8):1956–68.
124. Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, et al. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell. 2014;156(4):836–43.
125. Shapiro A, Lakey J, Ryan E. Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. N Engl J Med. 2000;343(4):230–8.
126. Kroon E, Martinson LA, Kadoya K, Bang AG, Kelly OG, Eliazer S, et al. Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. Nat Biotechnol. 2008;26(4):443–52.
127. Kelly OG, Chan MY, Martinson LA, Kadoya K, Ostertag TM, Ross KG, et al. Cell-surface markers for the isolation of pancreatic cell types derived from human embryonic stem cells. Nat Biotechnol. 2011;29(8):750–6.
128. Rezania A, Bruin JE, Xu J, Narayan K, Fox JK, O'Neil JJ, et al. Enrichment of human embryonic stem cell-derived NKX6.1-expressing pancreatic progenitor cells accelerates the maturation of insulin-secreting cells in vivo. Stem Cells. 2013;31(11):2432–42.
129. Rezania A, Bruin J, Riedel M, Mojibian M. Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. Diabetes. 2012;61(8):2016–29.
130. Capecchi MR. Altering the genome by homologous recombination. Science. 1989; 244(4910):1288–92.
131. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823–6.
132. Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, et al. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 1998;95(3):379–91.
133. Radzisheuskaya A, Le Bin CG, dos Santos RL, Theunissen TW, Castro LFC, Nichols J, et al. A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. Nat Cell Biol. 2013;15(6):579–90.
134. Gerrard L, Zhao D, Clark AJ, Cui W. Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency. Stem Cells. 2005;23(1):124–33.
135. Thyagarajan B, Liu Y, Shin S, Lakshmipathy U, Scheyhing K, Xue H, et al. Creation of engineered human embryonic stem cell lines using phiC31 integrase. Stem Cells. 2008;26(1):119–26.
136. Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. Nat Biotechnol. 2003;21(3):319–21.
137. Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol. 2009;27(9):851–7.
138. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011;29(8):731–4.
139. Rizzino A. Concise review: the Sox2-Oct4 connection: critical players in a much larger interdependent network integrated at multiple levels. Stem Cells. 2013;31(6):1033–9.
140. Brafman DA, Moya N, Allen-Soltero S, Fellner T, Robinson M, McMillen ZL, et al. Analysis of SOX2-expressing cell populations derived from human pluripotent stem cells. Stem Cell Rep. 2013;1(5):464–78.
141. Fischer Y, Ganic E, Ameri J, Xian X, Johannesson M, Semb H. NANOG reporter cell lines generated by gene targeting in human embryonic stem cells. PLoS One. 2010;5(9), e12533.
142. Zorn AM, Wells JM. Vertebrate endoderm development and organ formation. Annu Rev Cell Dev Biol. 2009;25(1):221–51.
143. Kearns NA, Genga RMJ, Enuameh MS, Garber M, Wolfe SA, Maehr R. Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. Development. 2014;141(1):219–23.

144. Balboa D, Weltner J, Eurola S, Trokovic R, Wartiovaara K, Otonkoski T. Conditionally stabilized dCas9 activator for controlling gene expression in human cell reprogramming and differentiation. Stem Cell Rep. 2015;5(3):448–59.
145. Wang P, Rodriguez RT, Wang J, Ghodasara A, Kim SK. Targeting SOX17 in human embryonic stem cells creates unique strategies for isolating and analyzing developing endoderm. Cell Stem Cell. 2011;8(3):335–46.
146. Jonsson J, Carlsson L, Edlund T, Edlund H. Insulin-promoter-factor 1 is required for pancreas development in mice. Nature. 1994;371(6498):606–9.
147. Offield M, Jetton T, Labosky P, Ray M, Stein R, Magnuson M, et al. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development. 1996;122(3):983–95.
148. Schwitzgebel VM, Mamin A, Brun T, Ritz-Laser B, Zaiko M, Maret A, et al. Agenesis of human pancreas due to decreased half-life of insulin promoter factor 1. J Clin Endocrinol Metab. 2003;88(9):4398–406.
149. Stoffers DA, Zinkin NT, Stanojevic V, Clarke WL, Habener JF. Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. Nat Genet. 2004;15(1):106–10.
150. Stoffers DA, Ferrer J, Clarke WL, Habener JF. Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. Nat Genet. 1997;17(2):138–9.
151. Nicolino M, Claiborn KC, Senée V, Boland A, Stoffers DA, Julier C. A novel hypomorphic PDX1 mutation responsible for permanent neonatal diabetes with subclinical exocrine deficiency. Diabetes. 2010;59(3):733–40.
152. Zhu Z, Verma N, González F, Shi Z-D, Huangfu D. A CRISPR/Cas-mediated selection-free knockin strategy in human embryonic stem cells. Stem Cell Rep. 2015;4(6):1–27.
153. Gu G, Dubauskaite J, Melton DA. Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. Development. 2002;129(10):2447–57.
154. Gradwohl G, Dierich A, LeMeur M. neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. Proc Natl Acad Sci U S A. 2000;97(4):1607–11.
155. Schwitzgebel VM, Scheel DW, Conners JR, Kalamaras J, Lee JE, Anderson DJ, et al. Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. Development. 2000;127(16):3533–42.
156. Apelqvist Å, Li H, Sommer L, Beatus P, Anderson DJ, Honjo T, et al. Notch signalling controls pancreatic cell differentiation. Nature. 1999;400(6747):877–81.
157. Pan FC, Wright C. Pancreas organogenesis: from bud to plexus to gland. Dev Dyn. 2011;240(3):530–65.
158. Piper K, Brickwood S, Turnpenny L. Beta cell differentiation during early human pancreas development. J Endocrinol. 2004;181(1):11–23.
159. Rubio-Cabezas O, Jensen JN, Hodgson MI, Codner E, Ellard S, Serup P, et al. Permanent neonatal diabetes and enteric anendocrinosis associated with biallelic mutations in NEUROG3. Diabetes. 2011;60(4):1349–53.
160. McGrath PS, Watson CL, Ingram C, Helmrath MA, Wells JM. The basic helix-loop-helix transcription factor NEUROG3 is required for development of the human endocrine pancreas. Diabetes. 2015;64(7):2497–505.
161. Teo AKK, Gupta MK, Doria A, Kulkarni RN. Dissecting diabetes/metabolic disease mechanisms using pluripotent stem cells and genome editing tools. Mol Metab. 2015;4(9):593–604.

Genome Editing in the Retina: A Case Study in CRISPR for a Patient-Specific Autosomal Dominant Retinitis Pigmentosa Model

Sally Justus, Andrew Zheng, Yi-Ting Tsai, Wen-Hsuan Wu, Chun-Wei Hsu, Wei-Pu Wu, Alexander G. Bassuk, Vinit B. Mahajan, and Stephen H. Tsang

Abstract The future of precision medicine, genome editing has gained momentum in the field of ophthalmology because of the eye's amenability to genetic interventions. The eye is an ideal target for gene therapy due to its accessibility, ease of noninvasive monitoring, significant compartmentalization, immunoprivileged status, optical transparency, and the presence of a contralateral control. One of the first gene therapy clinical trials was conducted in the eye for a severe form of early-onset retinal dystrophy called Leber congenital amaurosis, and it has encouraged further exploration of this technique as a viable treatment option for other inherited disorders across medical disciplines. This chapter highlights current ocular gene therapy approaches, clinical and preclinical experiments, and provides a case study of the bench-to-bedside personalized medicine approach taken for a novel and rare retinitis pigmentosa mutation.

Keywords Gene therapy • Gene repair • Gene supplementation • Retinitis pigmentosa • D190N mutation

S. Justus • A. Zheng • Y.-T. Tsai • W.-H. Wu • C.-W. Hsu • W.-P. Wu
Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, 635 West 165th Street, New York, NY 10032, USA

A.G. Bassuk
Department of Pediatrics, The University of Iowa Children's Hospital, 25 S. Grand Avenue, MedLabs Building 2040A, Iowa City, IA 52246, USA

V.B. Mahajan
Department of Ophthalmology, Iowa University Hospitals and Clinics, 200 Hawkins Drive, Iowa City, IA 52242, USA

S.H. Tsang, M.D., Ph.D. (✉)
Department of Ophthalmology, College of Physicians and Surgeons, Columbia University, 635 West 165th Street, New York, NY 10032, USA

Department of Pathology and Cell Biology, College of Physicians and Surgeons, Columbia University, 635 West 165th Street, New York, NY 10032, USA

Edward S. Harkness Eye Institute, New York-Presbyterian Hospital/Columbia University Medical Center, 635 West 165th Street, Box 112, New York, NY 10032, USA
e-mail: sht2@columbia.edu

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4_9

Retinitis Pigmentosa

Retinal dystrophies and retinitis pigmentosa (RP) are a heterogeneous group of inherited retinal dystrophies characterized by rod and cone photoreceptor cell death [1, 2]: 1 in 4000 individuals are affected, and more than 200 genes are implicated, leading to a spectrum of phenotypic manifestations and symptoms [2, 3]. Typically, patients experience night blindness in adolescence, tunnel vision in early adulthood, and blindness from a severely constricted visual field later in life (Fig. 1) [1, 4]. There are limited treatment options available. Currently, high-dose vitamin A palmitate, omega-3 supplements, and avoidance of vitamin E are the only widely accepted treatment recommendations, although they only slow progression and convey risks of birth defects and liver toxicity [1, 5, 6].

Scientists have turned to gene therapy in the hopes of developing an alternative, safe, lifelong treatment for RP that will halt cell degeneration indefinitely [7, 8]. Multiple gene therapy trials have centered on RP because many of the genes causing RP have already been cloned, and several well-characterized mouse models exist.

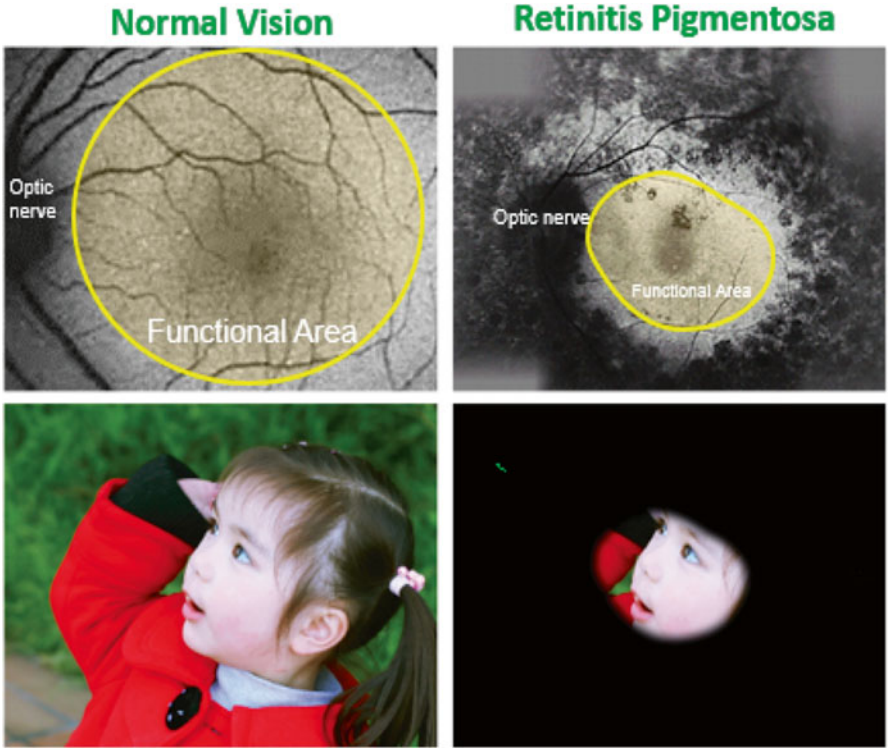

Fig. 1 Retinitis pigmentosa (RP) phenotypic and visual comparison. The *upper left* and *right boxes* show autofluorescence images of the retina. RP-affected retinae have fovea preservation amidst peripheral cell death, leading to tunnel vision as depicted in the *lower right box*

The interventional approach of the therapy is dependent on the inheritance pattern of the condition: recessive conditions can be amended through gene supplementation, whereas patients with dominant conditions must undergo gene correction.

Gene Supplementation

In recessive conditions, germline mutations in both alleles prevent production of a functional gene product. Supplementation of a wild-type copy of the gene can restore the normal phenotype [9, 10]. Gene supplementation distinguishes itself from gene correction in that the mutant alleles are left intact, relying on translation of the wild-type transgene to compensate for the inactive, absent, or defective protein. The most commonly used strategy for introducing the wild-type gene is viral transduction. Transgenes are packaged into viral vectors, which transduce the target cells, delivering the transgenes to the host in a cell-specific fashion [10]. Choice of vector allows for flexibility in therapy and determines the permanence of transgene expression, its pathogenicity, and the risk of insertional mutagenesis, among other factors [11]. A discussion of common vectors and their application in clinical and preclinical models is provided in the following section.

Viral Gene Therapy Strategies

Adenoviruses

The adenovirus is a non-enveloped, double-stranded DNA virus with broad tropism [9, 11]. The virus binds to cell-surface receptors, which internalize the DNA. The transgenes are not integrated into the genome, however, but exist as transcriptionally active episomes that are degraded on the timescale of days to weeks [12]. Adenovirus vector transgene expression is thus transient, and repeated vector administration is necessary for long-term clinical benefits. The status of the transgene as an episome can in other ways be advantageous, as its expression is immediate and independent of the cell reproductive state [11]. Concern regarding the pathogenicity of adenoviruses has led to the development of replication-deficient viruses, although there still remains the risk of recombination with wild-type adenoviruses within the host [13]. Encapsidated adenoviral mini-chromosomes (EAMs), which have had all viral genes removed, may address this limitation and are able to carry 30 kb DNA compared to only 8 kb in the average adenovirus vector [11]. The efficacy of EAMs was demonstrated by Kumar-Singh and Farber, who achieved photoreceptor functional rescue and transgene expression lasting 90 days following subretinal delivery (Fig. 2) of cyclic GMP phosphodiesterase cDNA to mice with retinal degeneration using EAM vectors [14, 15].

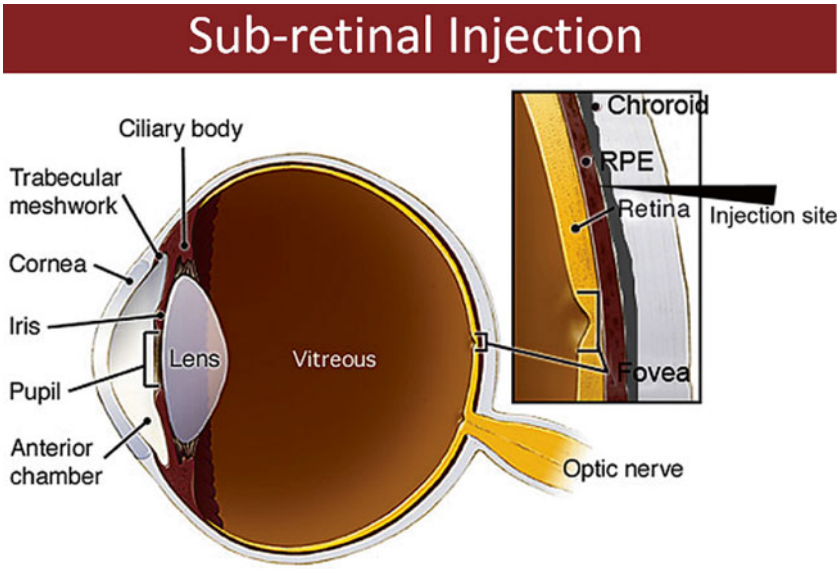

Fig. 2 Subretinal injections. A schematic diagram of the eye indicates the location of subretinal injections

Adeno-Associated Viruses

Adeno-associated virus (AAV) vectors are another popular delivery strategy. AAVs are non-enveloped, single-stranded DNA viruses that can infect mitotic and postmitotic cells, inserting transgenes into specific loci [16]. AAVs are not associated with any human infectious diseases, making them a safer option than adenoviruses, and they are also less likely to induce deleterious side effects from erroneous insertions [11]. However, they are comparatively limited in their cDNA carrying capacity (4.5 kb), and their expression occurs gradually rather than immediately.

Use of AAV vectors is common in ocular gene therapy. An AAV2/8 vector carrying a wild-type copy of the *PDE6a*^{nmf363} gene with a cell type-specific rhodopsin promoter was transduced in the retinas of *Pde6a*^{nmf363} mice, an RP model organism. A single injection preserved photoreceptor cells for at least 6 months compared to untreated mice, whose photoreceptors were undetectable by that time point [17]. Similar experiments were conducted in the autosomal recessive RP mouse model *Mfrp*^{rd6}/*Mfrp*^{rd6}, wherein long-term visual function preservation was achieved through AAV2/8-hu*MFRP* infection [18]. These findings were corroborated in in vitro patient-derived, induced pluripotent stem (iPS)-retinal pigment epithelium (RPE) cell lines, in which AAV infection restored the cell pigmentation, transepithelial resistance, actin structural organization, and apical microvilli. Several clinical trials based on AAV vector therapy were conducted for Leber congenital amaurosis, a

severe form of retinitis pigmentosa caused by a loss-of-function mutation in the 65-kDa protein of the retinal pigment epithelium (RPE65) [19–21]. Researchers were able to successfully express the transgene by way of a rAAV-RPE65 vector and temporarily halt the photoreceptor degeneration, although cell death and functional impairment resumed after 3 years. Research to determine the ideal delivery time and amount is underway to address this issue.

RNA Viruses

Many RNA viruses, including retroviruses, integrate into the host genome at random locations and risk inducing insertional mutagenesis. They have a DNA-shuttling capacity of 8 kb and are limited to transducing mitotic cells because they are unable to enter the nucleus of nondividing cells. Lentiviruses, on the other hand, are a subclass of RNA retroviruses that can form complexes with the nuclear envelope of nondividing cells, meaning they can infect mitotic and post-mitotic cells alike [22].

Lentiviruses have undergone significant modifications to enhance their safety for clinical use. Pseudo-typed viruses, where only transgenes, not viral genes, integrate into the host, have been developed to avoid activation of T-cell lymphocytes, which trigger inflammation and destruction of vector-containing cells [22]. Concerns regarding the clinical application of lentiviruses, which could theoretically generate live HIV through three recombinations, have motivated further efforts to develop new constructs with decreased risks.

Lentiviral vectors have been successfully used in mouse models of RP to halt photoreceptor degeneration. Lentiviral-mediated gene transfer of *Opsin*::*Pde6b* in *Pde6b*^{H620Q} homozygous mice delayed photoreceptor death [23]. In a follow-up study using the same preclinical model as in Davis et al., bipartite vectors that expressed both the wild-type *Pde6b* gene and small hairpin RNA (shRNA) encoding guanlyate cyclase (GUCY) facilitated significantly greater photoreceptor survival compared to the monopartite vectors. Normally in the phototransduction cascade, the *Pde6b* protein is activated when photons impinge on rod outer segments, and it reduces free cGMP concentration via hydrolysis; this in turn lowers the excitation state of the photoreceptor and halts the release of neurotransmitter from the rod. Contrastingly, GUCY is the opposing counterpart of *Pde6b* and stimulates cGMP production from guanosine-5′-triphosphate (GTP). In darkness, *Pde6b* is inactivated, leading to high cGMP and calcium levels. For homozygous mutants such as *Pde6b*$^{H620Q/H620Q}$, *Pde6b* is inactive, and perpetually elevated levels of cGMP and calcium eventually cause cell death. Thus, the halt in photoreceptor degeneration observed in this study was plausibly obtained by expressing the wild-type transgene while simultaneously lowering calcium levels. This observation suggests that combined gene therapies may be more efficacious than single therapies alone [24].

Gene Correction and Repair

In dominant conditions, one or more alleles carry a dominant negative or gain-of-function mutation that cannot be remedied by introducing the wild-type copy of the gene. Instead, the mutated allele(s) must be silenced or removed and, in some cases, the wild-type gene must be inserted in its place. This insertion poses the challenge of a two-step process, wherein the mutant allele is removed or silenced and the functional gene is introduced. Several approaches to remove or override the mutated gene have been developed, including ribozyme therapy, antisense gene therapy, antiapoptosis therapy, and the newest technologies: zinc-finger nucleases (ZFN), transcription activator-like effector nucleases (TALEN), and clustered regularly interspaced short palindromic repeats (CRISPR).

Ribozyme Interference

Ribozymes cleave specific messenger RNA (mRNA) sequences by hydrolyzing trinucleotide motifs, preventing protein translation [25]. They are synthesized in situ from vectors and are a useful technique for gene correction by disrupting a dominant mutant gene protein product [26]. They can even be used in gene correction strategies. In a mutation-independent fashion, the ribozyme can be designed to cleave both the wild-type and mutant mRNA transcripts; then, a functional gene copy not targetable by the ribozyme can be introduced into the cell [11]. Lewin et al. successfully executed this technique and slowed photoreceptor degeneration for 3 months in a rat model of autosomal dominant RP using ribozymes to silence the mutant mRNAs produced by a proline to histidine substitution at codon 23 of the rhodopsin gene (P23H) [27].

Antisense Gene Therapy

With many parallels to ribozyme interference, antisense gene therapy uses short, synthetic DNA sequences that are complementary to targeted mRNA transcripts to form DNA–RNA heteroduplexes that prevent translation [11]. RNase H degrades the mRNA in the duplex, releasing the oligodeoxynucleotide (ODN) to bind to another complementary mRNA strand. In a study of ischemia-induced proliferative retinopathies, antisense ODNs complementary to the mRNA of the neovascularization gene product, vascular endothelial growth factor (VEGF), were administered before the onset of proliferative retinopathy [28]. The VEGF protein concentration was reduced by 40–66 % and blood vessel growth by 25–31 %, demonstrating the efficacy of this technique.

Antiapoptosis Gene Therapy

Programmed cell death is regulated by apoptotic proteins such as *p53*. Researchers have modulated the genes that control these apoptotic proteins to prevent photoreceptor death in a mutation-independent manner. For example, viral transfer of the antiapoptotic gene *bcl*-2 to the retina was shown to delay degeneration in various mouse models [29]. However, transfer to retinal ganglion cells had the opposite effect, leading to increased apoptosis, suggesting that the effects of these genes may be cell type specific [30]. This technique is promising and merits further research and exploration.

Recent Advances: ZFN, TALEN, and CRISPR

Scientists have studied naturally occurring DNA-repair mechanisms to inspire new approaches for targeted gene therapy. A common, natural DNA-editing scheme occurs in response to double-strand breaks (DSBs) in DNA caused unintentionally by external forces, such as irradiation, or intentionally for the purposes of recombination [31, 32]. DSBs activate two main repair strategies. The first is non-homologous end-joining (NHEJ), wherein the ends of the cleaved DNA strands are simply religated together without regard for sequence accuracy [31, 33]. This error-prone process causes small insertions or deletions (indels) that often lead to nonfunctional gene products. The second repair strategy is homology-directed repair (HDR), in which the broken strands are paired with an unbroken sister chromatid or homologous chromosome, which serves as a template for accurately reforming the DNA sequence at the break site. Exploiting these repair mechanisms for the purposes of targeted gene therapy was delayed by the challenge of finding a reliable means of inducing DSBs at specific loci. Protein nucleases were recognized as having the most potential for addressing this challenge, and the zinc-finger nuclease served as the first generation of nuclease-based gene therapies.

Zinc-Finger Nucleases

Synthetic zinc-finger nucleases (ZFN) are made by fusing zinc-finger-binding domains to the catalytic unit of the type IIS restriction enzyme, *Fok*I [31, 32]. Each finger contains roughly 30 amino acids, 4 of which (Cys_2His_2) coordinate a zinc atom. One zinc finger interacts with three base pairs, and at least three consecutive fingers are required for adequate binding. The catalytic domain of *Fok*I must dimerize to cleave the DNA, necessitating two ZFNs per site.

ZFN technology has been successfully applied in the retina. Greenwald et al. [34] designed ZFNs to target human embryonic retinoblast cell lines and found ZFN-induced homologous recombination rates at the human rhodopsin gene to be

as high as 17 % compared with endogenous homologous recombination [34]. Another study examined the applications of ZFNs for Usher syndrome, a form of RP that leads to deafness and blindness. Although five USH1 genes (USH1B-G) have been identified, Overlack and colleagues focused on the p.R31X mutation in murine *Ush1c*, for which they designed customized ZFNs [35]. Their data confirmed the feasibility of site-specific gene correction following treatment with ZFNs and introduction of rescue DNA.

A drawback of ZFNs is that not all variations of three base pairs can be targeted because zinc-finger domains for certain combinations of nucleotides have not yet been engineered. Reliable specificity and affinity are difficult to achieve as well, and neighboring domains in the protein may affect the specificity of binding [31]. The unreliability of ZFNs has motivated scientists to search for alternatives.

Transcription Activator-Like Effector Nucleases

Plant-pathogenic bacteria in the genus *Xanthomonas* integrate into the host genome and promote infection using genes that encode for transcription activator-like endonucleases (TALs) [31, 32]. The TAL proteins are DNA-binding domains with tandem repeats of 34 amino acids that are highly conserved except at residues 12 and 13, which vary based on the identity of the nucleotide of the target sequence. TALs were fused with the catalytic domain of *Fok*I to recapitulate ZFNs, and the complex was called TALEN (TAL effector nuclease). Similar to ZFN, TALEN utilizes *Fok*I and thus must dimerize to cleave the DNA. However, they differ in that there is one TAL per nucleotide, which greatly simplifies the recognition code compared to ZFNs and confers greater binding efficiency. The disadvantage of the one-to-one ratio of base pair to TALs is that long arrays of TALs (at least 12) are required to achieve adequate binding affinity and specificity.

TALENs were used to create an X-linked RP 2 (RP2) zebrafish model that recapitulates the retinal degeneration typical of affected humans by disrupting the RP2 gene [36]. In another application of TALEN, iPS cell lines derived from best vitelliform macular dystrophy (BVMD) patients were corrected using TALEN by inducing NHEJs following DSB and disrupting the vitelliform macular degeneration 2 (VMD2) gene product [37]. In yet another experiment, ocular abnormalities caused by mutations in the *Crb1* gene were corrected by HDR following TALEN-induced DSB in murine models. In this study, mouse embryos with the $Crb1^{rd8}$ mutation were coinjected with TALENs targeting the *Crb1* gene loci and single-stranded, complementary oligonucleotides to correct the allele, with only slight nucleotide alterations in the strand to avoid targeting by TALEN. Twenty-seven percent of offspring showed HDR, although founder mosaicism was present among them. Treated mice showed the normal retinal phenotype, but the untreated mice exhibited retinal dysplasia [38].

Clustered Regularly Interspaced Short Palindromic Repeats

Most recently, research in bacteria and archaea immunity strategies has led to the creation of the revolutionary CRISPR/Cas9 technology. It was discovered that specific regions of the bacterial genome are reserved for the incorporation of invading viral genes, which are broken into segments and integrated within the bacterial genome in between clustered regularly interspaced short palindromic repeats (CRISPR) [39]. The incorporated viral genes are then expressed as CRISPR-derived RNA (crRNA), which complexes with the CRIPSR-associated (Cas) nuclease and a trans-activating crRNA (tracrRNA), which is involved in crRNA biogenesis and maturation in the type II CRISPR/Cas system [33, 40]. The crRNA–tracrRNA complex acts as a guide for the Cas protein and directs it to specific sequences on the target strand, where it induces DSBs.

In 2012, Jinek and colleagues fused the crRNA and tracrRNA to create chimeric guide RNA (gRNA), which they demonstrated to be equally as effective as the crRNA–tracrRNA complex in guiding the Cas9 protein to cleave DNA and is customizable to target any number of sites [41]. The guide RNA has 20 amino acids at its 5′-end that determine its target specificity and are alterable. The 3′-end contains an invariable region required to complex with the Cas9 protein. To cleave DNA, the target sequence must contain a protospacer-adjacent motif (PAM) directly downstream of the desired cleavage site [42]. There is no analogue to PAM in the gRNA. The *Streptococcus pyogene*s Cas9 protein recognizes the PAM, NGG, where N is variable and G is the guanosine nucleotide [43]. Delivering the Cas9 protein can be accomplished through a helper plasmid that carries the coding sequence and a promoter, which may be fused with a nuclear localization signal [44]. Introducing the guide RNA may be more difficult but can be achieved by expressing the gRNA from a plasmid or transgene in vivo in conjunction with RNA polymerase III promoters [45].

The advantage of CRISPR/Cas9 over TALEN is that it has a much higher targeting efficiency, and CRISPR/Cas9 has the ability to cleave regardless of the target sequence DNA methylation status [43]. CRISPR/Cas9 specificity is arguably more precise, especially with the development of nickases to minimize off-targeting effects. A mutant Cas9, called a nickase, cleaves only one strand of the target DNA sequence. Nickases can still create a DSB by inducing two single-strand breaks (nicks) close to one another on opposite DNA strands [46, 47]. Requiring these nicks to be closely linked drastically reduces off-targeting effects and enhances targeting specificity.

CRISPR/Cas9 is a 3-year-old technology that is rapidly evolving, but in a short time it has already overtaken other nuclease-based strategies and has been applied in numerous diverse model organisms (review by Sander and Young) [39]. There is every indication that its applications will continue to expand, becoming increasingly far reaching.

Case Study in the Bench-to-Bedside Approach for the Autosomal Dominant RP D190N Mutation

A family presented with an RP-inducing rhodopsin mutation (Asp190Asn, or D190N) inherited in an autosomal dominant fashion. Rhodopsin (RHO) is a G protein-coupled receptor located in the rod outer segment discs and is responsible for initiating the phototransduction cascade. Mutations in *RHO* account for roughly 54 % of autosomal dominant RP (adRP) conditions [48]. Normally, photons induce the isomerization of 11-*cis* retinal to all-*trans* retinal on the RHO chromophore. This process malfunctions in D190N individuals, and the mutation gradually leads to cell death in a manner that has not yet been determined. This finding motivated the development of a mouse model that recapitulates the phenotype and natural history of the disease.

Heterozygous D190N knock-in mice were produced as described by Sancho and colleagues, and homozygous mutants were bred from the heterozygotes [48]. Electrophysiology, histology, angiography, imaging results, and functional measurements were shown to faithfully recapitulate the human disease. Mouse eyes were then transduced with dual *AAV8::Cas9* and *AAV8::gRNA* vectors injected subretinally (Fig. 3). Cone cells, the death of which is the primary cause of blindness, were preserved in CRISPR/Cas9-mediated rescue (Fig. 4). Electroretinogram data on a D190N heterozygote demonstrated functional rescue as well, with treated mice having greater b-wave amplitudes compared to controls (Fig. 5). These preliminary data suggest that CRISPR/Cas9 may be a viable option for correcting this novel *RHO* mutation in the patients.

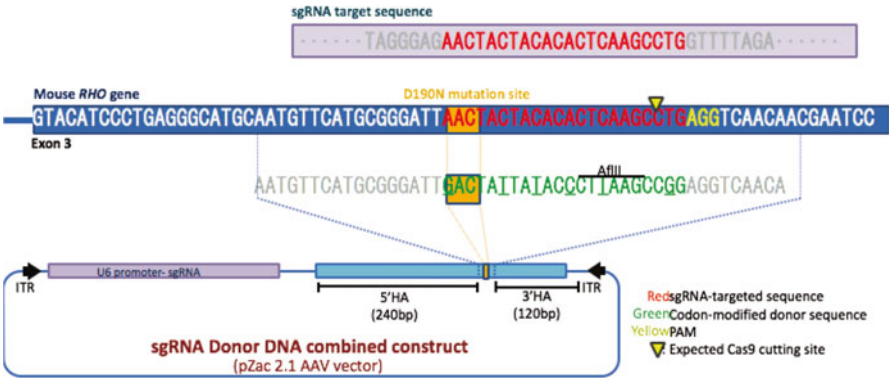

Fig. 3 Adeno-associated virus (AAV) construct carrying Cas9 and repair templates. The pZac AAV vector was used to transduce double-dominant *Rho* (Rho^{D190N}/Rho^{D190N}) models

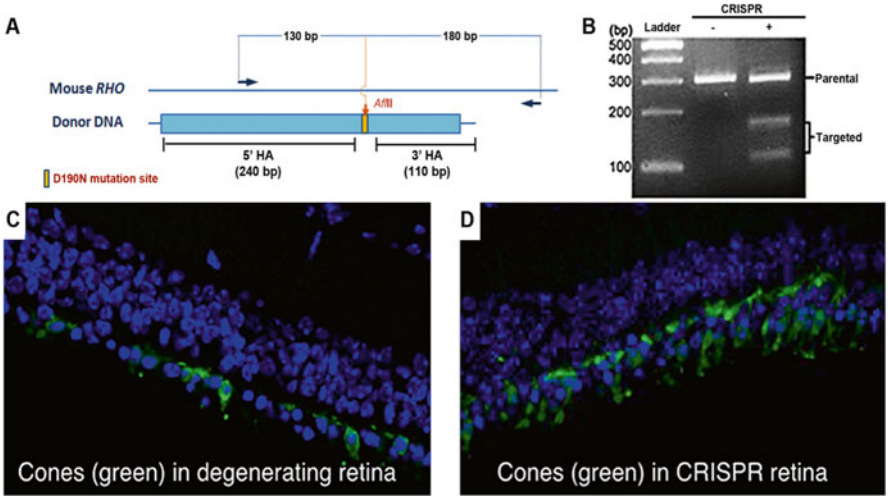

Fig. 4 AAV-mediated repair of a double-dominant *Rho* model. Verification of guide RNA (gRNA) targeting ability by *Afl*II enzyme digestion. (**a**) Mouse *Rho* genomic sequence was replaced with a codon-modified donor sequence containing an additional *Afl*II site (*orange arrow*), marking cells that underwent homologous recombination; *dark blue arrows* represent primer pairs for PCR amplification. (**b**) Amplicons generated from cells undergoing recombination are 130 and 180 bp; the parental sequence is a single 310-bp band. (**c**) Retinal section of a 3-month-old homozygote Rho^{D190N}/Rho^{D190N} eye transduced by the *AAV8::Cas9* vector reveals only a few cones (*green*). (**d**) Rho^{D190N}/Rho^{D190N} eye co-transduced with donor template and Cas9 vectors indicates cone rescue (*green*)

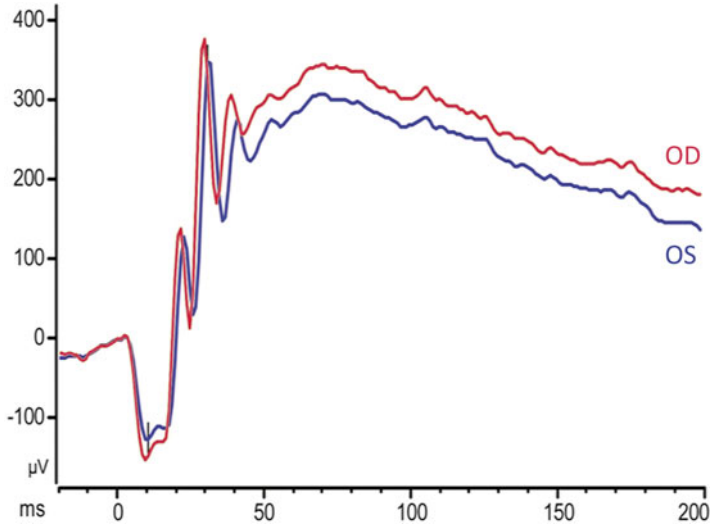

Fig. 5 Functional data. Functional rescue is detectable in the experimental right eye (OD) compared with the control left eye (OS), as indicated by the higher b-wave amplitude on the electroretinogram (ERG) recording

Conclusions

The past few years have seen significant advances in gene editing. Targeting efficiency and binding specificity have been enhanced significantly, and the combined use of viral vectors alongside DNA-editing technologies such as TALEN and CRISPR has improved site-specific gene targeting in a way that has never been achieved before. With the latest developments in genome editing, the future of precision medicine has never been brighter.

Acknowledgments Funding/Support: V.B.M. is supported by the National Institutes of Health [K08EY020530, R01EY016822, R01EY024665, R01EY025225, R01EY024698 and R21AG050437], Doris Duke Charitable Foundation Grant #2013103, and Research to Prevent Blindness, New York, NY. G.V. is supported by the NIH [T32GM007337]. The Barbara & Donald Jonas Laboratory of Regenerative Medicine is supported by the National Institute of Health Core [5P30EY019007], National Cancer Institute Core [5P30CA013696], unrestricted funds from Research to Prevent Blindness (RPB), New York, NY, USA. S.H.T. is a member of the RD-CURE Consortium and is supported by the Tistou and Charlotte Kerstan Foundation, the National Institute of Health [R01EY018213], the Research to Prevent Blindness Physician-Scientist Award, Association for Research in Vision and Ophthalmology (ARVO) Foundation, Macula Society, the Schneeweiss Stem Cell Fund, New York State [C029572], the Foundation Fighting Blindness New York Regional Research Center Grant [C-NY05-0705-0312], the Joel Hoffman Fund, the Professor Gertrude Rothschild Stem Cell Foundation, and the Gebroe Family Foundation. V.B.M. is supported by the National Institutes of Health [K08EY020530, R01EY016822], Doris Duke Charitable Foundation Grant #2013103, and Research to Prevent Blindness, New York, NY. G.V. is supported by the NIH [T32GM007337].

References

1. Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. Lancet. 2006;368:1795–809.
2. Hamel C. Retinitis pigmentosa. Orphanet J Rare Dis. 2006;1:40.
3. Parmeggiani F. Clinics, epidemiology and genetics of retinitis pigmentosa. Curr Genomics. 2011;12:236–7.
4. Fahim AT, Daiger SP, Weleber RG. Retinitis pigmentosa overview. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJH, Bird TD, Fong CT, Mefford HC, Smith RJH, Stephens K, editors. GeneReviews(R). Seattle: University of Washington; 1993.
5. Lin MK, Tsai YTT, Tsang SH. Emerging treatments for retinitis pigmentosa: genes and stem cells, as well as new electronic and medical therapies, are gaining ground. Retin Physician. 2015;12:52–70.
6. Berson EL, Rosner B, Sandberg MA, Hayes KC, Nicholson BW, Weigel-DiFranco C, Willett W. A randomized trial of vitamin A and vitamin E supplementation for retinitis pigmentosa. Arch Ophthalmol. 1993;111:761–72.
7. Wu WH, Tsai YT, Justus S, Lee T, Zhang L, Lin CS, Bassuk AG, Mahajan VB, Tsang SH. CRISPR repair reveals causative mutation in a preclinical model of retinitis pigmentosa. Mol Ther. 2016. In Press.
8. Bassuk AG, Zheng A, Li Y, Tsang SH, Mahajan VB. Precision Medicine: Genetic Repair of Retinitis Pigmentosa in Patient-Derived Stem Cells. Scientific reports. 2016;6(19969).
9. Dalkara D, Sahel J-AA. Gene therapy for inherited retinal degenerations. C R Biol. 2014;337:185–92.

10. Boye SE, Boye SL, Lewin AS, Hauswirth WW. A comprehensive review of retinal gene therapy. Mol Ther. 2013;21:509–19.
11. Chaum E, Hatton MP. Gene therapy for genetic and acquired retinal diseases. Surv Ophthalmol. 2002;47:449–69.
12. McClements ME, MacLaren RE. Gene therapy for retinal disease. Transl Res. 2013;161: 241–54.
13. Nash BM, Wright DC, Grigg JR, Bennetts B, Jamieson RV. Retinal dystrophies, genomic applications in diagnosis and prospects for therapy. Transl Pediatr. 2015;4:139–63.
14. Wert KJ, Skeie JM, Davis RJ, Tsang SH, Mahajan VB. Subretinal injection of gene therapy vectors and stem cells in the perinatal mouse eye. J Vis Exp 2012;69:
15. Kumar-Singh R, Farber DB. Encapsidated adenovirus mini-chromosome-mediated delivery of genes to the retina: application to the rescue of photoreceptor degeneration. Hum Mol Genet. 1998;7:1893–900.
16. Kotterman MA, Schaffer DV. Engineering adeno-associated viruses for clinical gene therapy. Nat Rev Genet. 2014;15:445–51.
17. Wert KJ, Davis RJ, Sancho-Pelluz J, Nishina PM, Tsang SH. Gene therapy provides long-term visual function in a pre-clinical model of retinitis pigmentosa. Hum Mol Genet. 2013;22:558–67.
18. Li Y, Wu W-H, Hsu C-W, Nguyen HV, Tsai Y-T, Chan L, Nagasaki T, Maumenee IH, Yannuzzi LA, Hoang QV, Hua H, Egli D, Tsang SH. Gene therapy in patient-specific stem cell lines and a preclinical model of retinitis pigmentosa with membrane frizzled-related protein defects. Mol Ther. 2014;22:1688–97.
19. Bainbridge JW, Smith AJ, Barker SS, Robbie S, Henderson R, Balaggan K, Viswanathan A, Holder GE, Stockman A, Tyler N, Petersen-Jones S, Bhattacharya SS, Thrasher AJ, Fitzke FW, Carter BJ, Rubin GS, Moore AT, Ali RR. Effect of gene therapy on visual function in Leber's congenital amaurosis. N Engl J Med. 2008;358:2231–9.
20. Cideciyan AV, Aleman TS, Boye SL, Schwartz SB, Kaushal S, Roman AJ, Pang JJ, Sumaroka A, Windsor EA, Wilson JM, Flotte TR, Fishman GA, Heon E, Stone EM, Byrne BJ, Jacobson SG, Hauswirth WW. Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. Proc Natl Acad Sci U S A. 2008;105:15112–7.
21. Maguire AM, Simonelli F, Pierce EA, Pugh Jr EN, Mingozzi F, Bennicelli J, Banfi S, Marshall KA, Testa F, Surace EM, Rossi S, Lyubarsky A, Arruda VR, Konkle B, Stone E, Sun J, Jacobs J, Dell'Osso L, Hertle R, Ma JX, Redmond TM, Zhu X, Hauck B, Zelenaia O, Shindler KS, Maguire MG, Wright JF, Volpe NJ, McDonnell JW, Auricchio A, High KA, Bennett J. Safety and efficacy of gene transfer for Leber's congenital amaurosis. N Engl J Med. 2008;358:2240–8.
22. Balaggan KS, Ali RR. Ocular gene delivery using lentiviral vectors. Gene Ther. 2012;19:145–53.
23. Davis RJ, Hsu CW, Tsai YT, Wert KJ, Sancho-Pelluz J, Lin CS, Tsang SH. Therapeutic margins in a novel preclinical model of retinitis pigmentosa. J Neurosci. 2013;33:13475–83.
24. Tosi J, Sancho-Pelluz J, Davis RJ, Hsu CW, Wolpert KV, Sengillo JD, Lin C-SS, Tsang SH. Lentivirus-mediated expression of cDNA and shRNA slows degeneration in retinitis pigmentosa. Exp Biol Med (Maywood). 2011;236:1211–7.
25. Lewin AS, Hauswirth WW. Ribozyme gene therapy: applications for molecular medicine. Trends Mol Med. 2001;7:221–8.
26. Sioud M (2006) Ribozymes and siRnas: from structure to preclinical applications. In: Handbook of experimental pharmacology. Springer, Berlin, pp 223–242.
27. Gorbatyuk M, Justilien V, Liu J, Hauswirth WW, Lewin AS. Preservation of photoreceptor morphology and function in P23H rats using an allele independent ribozyme. Exp Eye Res. 2007;84:44–52.
28. Robinson GS, Pierce EA, Rook SL. Oligodeoxynucleotides inhibit retinal neovascularization in a murine model of proliferative retinopathy. Proc Natl Acad Sci U S A. 1996;93:4851–6.
29. Tsang SH, Chen J, Kjeldbye H, Li WS, Simon MI, Gouras P, Goff SP. Retarding photoreceptor degeneration in Pdegtm1/Pdegtml mice by an apoptosis suppressor gene. Invest Ophthalmol Vis Sci. 1997;38:943–50.
30. Simon PD, Vorwerk CK, Mansukani SS, Chen SJ, Wilson JM, Zurakowski D, Bennett J, Dreyer EB. bcl-2 gene therapy exacerbates excitotoxicity. Hum Gene Ther. 1999;10:1715–20.

31. Gilles AF, Averof M. Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. EvoDevo. 2014;5:43.
32. Carroll D. Genome engineering with targetable nucleases. Annu Rev Biochem. 2014;83:409–39.
33. Cong L, Ran AF, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339:819–23.
34. Greenwald DL, Cashman SM, Kumar-Singh R. Engineered zinc finger nuclease-mediated homologous recombination of the human rhodopsin gene. Invest Ophthalmol Vis Sci. 2010;51:6374–80.
35. Overlack N, Goldmann T, Wolfrum U, Nagel-Wolfrum K. Gene repair of an Usher syndrome causing mutation by zinc-finger nuclease mediated homologous recombination. Invest Ophthalmol Vis Sci. 2012;53:4140–6.
36. Liu F, Chen J, Yu S, Raghupathy RK. Knockout of RP2 decreases GRK1 and rod transducin subunits and leads to photoreceptor degeneration in zebrafish. Hum Mol Genet. 2015;24: 4648–59.
37. Sharma R, Rajan V, Wan Q. Engineering isogenic best vitelliform macular dystrophy patient iPS cell lines using TALEN technology. Invest Ophthalmol Vis Sci. 2014;55:3987.
38. Low BE, Krebs MP, Joung JK, Tsai SQ. Correction of the Crb1rd8 allele and retinal phenotype in C57BL/6N mice via TALEN-mediated homology-directed repair. Invest Ophthalmol Vis Sci. 2014;55:387–95.
39. Sander JD, Joung KJ. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32:347–55.
40. Senís E, Fatouros C, Große S, Wiedtke E, Niopek D, Mueller AK, Börner K, Grimm D. CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. Biotechnol J. 2014;9:1402–12.
41. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816–21.
42. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346:1258096.
43. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8:2281–308.
44. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157:1262–78.
45. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep. 2014;4:5400.
46. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X, Skarnes WC. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Methods. 2014;11:399–402.
47. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013;154:1380–9.
48. Sancho-Pelluz J, Tosi J, Hsu C-WW, Lee F, Wolpert K, Tabacaru MR, Greenberg JP, Tsang SH, Lin C-SS. Mice with a D190N mutation in the gene encoding rhodopsin: a model for human autosomal-dominant retinitis pigmentosa. Mol Med. 2012;18:549–55.

Index

K. Turksen (ed.), *Genome Editing*, DOI 10.1007/978-3-319-34148-4