

*Review*

## Minimizing off-Target Mutagenesis Risks Caused by Programmable Nucleases

Kentaro Ishida <sup>1,†</sup>, Peter Gee <sup>1,†</sup> and Akitsu Hotta <sup>1,2,\*</sup>

<sup>1</sup> Center for iPS Cell Research and Application (CiRA), Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan; E-Mails: kentaro.ishida@cira.kyoto-u.ac.jp (K.I.); peter.gee@cira.kyoto-u.ac.jp (P.G.)

<sup>2</sup> Institute for Integrated Cell Material Sciences (iCeMS), Kyoto University, Yoshida Ushinomiya-cho, Sakyo-ku, Kyoto 606-8507, Japan

† These authors contributed equally to this work.

\* Author to whom correspondence should be addressed; E-Mail: akitsu.hotta@cira.kyoto-u.ac.jp; Tel.: +81-75-366-7056; Fax: +81-75-366-7074.

Academic Editor: Izuho Hatada

*Received: 26 August 2015 / Accepted: 8 October 2015 / Published: 16 October 2015*

---

**Abstract:** Programmable nucleases, such as zinc finger nucleases (ZFNs), transcription activator like effector nucleases (TALENs), and clustered regularly interspersed short palindromic repeats associated protein-9 (CRISPR-Cas9), hold tremendous potential for applications in the clinical setting to treat genetic diseases or prevent infectious diseases. However, because the accuracy of DNA recognition by these nucleases is not always perfect, off-target mutagenesis may result in undesirable adverse events in treated patients such as cellular toxicity or tumorigenesis. Therefore, designing nucleases and analyzing their activity must be carefully evaluated to minimize off-target mutagenesis. Furthermore, rigorous genomic testing will be important to ensure the integrity of nuclease modified cells. In this review, we provide an overview of available nuclease designing platforms, nuclease engineering approaches to minimize off-target activity, and methods to evaluate both on- and off-target cleavage of CRISPR-Cas9.

**Keywords:** CRISPR Cas9; genome editing; mutagenesis; off-target effect

---

## 1. Introduction

Genome engineering technologies can be designed to recognize and induce double stranded breaks (DSBs) at specific DNA sequences, which are mainly repaired by two host DNA repair pathways. The nonhomologous end joining (NHEJ) pathway, which induces insertions or deletions (indels) at the cleavage site, can disrupt gene open reading frames, and the homologous recombination (HR) pathway which, in the presence of a complementary DNA sequence, exchanges damaged DNA with a homologous sequence. These technologies have allowed scientists to investigate the functional roles of genes in cells and live animals via gene knockout or knock-in approaches. In recent years, three types of endonucleases, namely ZFNs (zinc finger nucleases), TALENs (transcription activator like effector nucleases), and the CRISPR (clustered regularly interspersed short palindromic regions) associated 9 (Cas9) system have been predominantly utilized for gene editing. However, due to technical challenges involved in constructing specific ZFNs and TALENs, the CRISPR Cas9 system has emerged in recent years as the nuclease of choice because of its ease of construction, by simply designing RNA with a 20 base pair complementary sequence to a targeted region in the genome.

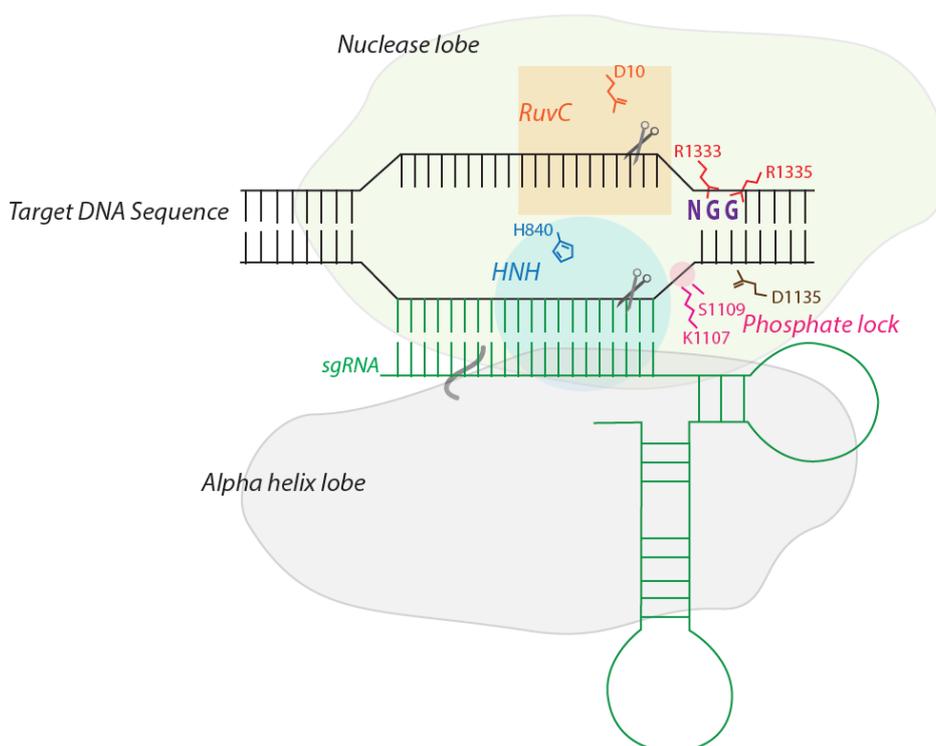
The CRISPR Cas9 system was first identified as an adaptive immune host defense mechanism in bacteria against invading pathogens such as bacteriophages [1–3]. The system functions by incorporating DNA fragments of pathogens from past infections into specialized regions of the bacterial genome that are transcribed into crRNA (CRISPR RNA) and bind in tandem with tracrRNA (transactivating CRISPR RNA) to Cas9 endonucleases in a complex that acts as a sentinel for potential threats with complementary DNA sequences [1]. Successful targeting induces DSBs in the foreign DNA, thereby eliminating the threat.

Since its discovery in 2012, the CRISPR Cas9 system has been harnessed into a powerful genome editing tool for scientists to modify the genomes of cell lines [4–7] and even living organisms (*C. elegans* [8,9], zebrafish [10], mice [11,12], rat [13,14], rabbits [15] and monkey [16]), with relative ease, compared to TALENs and ZFNs. The CRISPR Cas9 system also has tremendous therapeutic potential [17] and has been shown to correct genetic mutations [18] such as  $\beta$ -thalassemia [19], cystic fibrosis [20], hemophilia A [21], cataracts [22], hereditary tyrosinemia type I [23], and Duchenne muscular dystrophy [24]. Furthermore, it has been implemented for host defense in human cells as a tool to eradicate chronic virus infections including hepatitis B [25,26] and HIV [27–29].

However, given the early success of Cas9 thus far, accumulating lines of evidence point to a variable level of nonspecific activity at off-target sequences that has caused concern, particularly in the event that these nucleases are used in the clinical setting because the effects could be permanent and detrimental for cellular function, resulting in possible cellular toxicity or even tumorigenicity. The phenomena of non-specific cleavage by nucleases leading to cellular toxicity is not unique to CRISPR Cas9 as it has been also reported for ZFNs and TALENs [30–34], but the parameters directing RNA mediated sequence recognition and cleavage by Cas9 are unique because the targeting mechanism is determined by a complementary RNA sequence [35]. In this review, we will highlight strategies for minimizing off-target cleavage by using available online tools for targeting design, modified Cas9 enzymes to attenuate off-target activity, and cover methods to assess off-target cleavage activity.

## 2. The Structure of Cas9: How it Recognizes and Cleaves DNA

The crystal structure of Cas9 derived from *Streptococcus pyogenes* (Sp) has been resolved by several groups, and has helped to elucidate the atomic-level interactions between a complex of Cas9, crRNA, tracrRNA and target DNA, and the dynamic conformational changes the protein undergoes during substrate binding and cleavage [36–38]. Essentially, a Cas9 protein consists of two regions shaped like large lobes, one of which harbors two nuclease activities, RuvC and HNH, and the other, an  $\alpha$  helix (Figure 1). The main catalytic residues responsible for cleavage activity in RuvC and HNH are D10 and H840, respectively (Figure 1). Alone, Cas9 exists in an inactive state that undergoes conformational change upon guide RNA (sgRNA) binding, causing its two lobes to pull apart from one another and create a channel that is large enough to accommodate a target double strand DNA and sgRNA (Figure 1) [36]. Amino acids lining the channel are positively charged and interact with the phosphate backbone of the nucleic acids.



**Figure 1.** *Streptococcus pyogenes* (Sp) Cas9 recognition of target DNA. Sp Cas9 consists of two lobes, one containing the HNH and RuvC nuclease domains, and an  $\alpha$  helix lobe. Arginine residues (R1333 and R1335) of Sp Cas9 recognize the dinucleotide guanine bases of the “NGG” protospacer adjacent motif (PAM). Then, serine (S1109) and lysine (K1107) form a phosphate lock by interacting with a phosphate residue adjacent to the PAM site, which triggers the melting of the double stranded DNA so that sgRNA (green) can form Watson–Crick pairing with a complementary DNA strand. D10 and H840 are the main catalytic residues in the HNH and RuvC nuclease domains, respectively. Aspartic acid at 1135 (D1135) interacts with the third PAM guanine residue through a water molecule, and a modification of the residue to glutamic acid (D1135E) has been shown to improve the recognition of PAM sequence and target specificity.

If a trinucleotide “NGG” PAM (protospacer adjacent motif) exists adjacent to the 3' end of the target sequence on the noncomplementary strand, then a PAM interacting domain (R1333 and R1335) binds to the “GG” dinucleotide sequence (Figure 1) [39]. D1335 is another residue that interacts with the third PAM guanine via a water molecule. Opening of the dsDNA from the PAM site exposes a phosphate at the adjacent position, which interacts with lysine and serine residues (K1107 and S1109), termed a phosphate lock (Figure 1) [39]. This interaction triggers the unzipping of the double stranded DNA so that crRNA can form Watson–Crick pairing with the complementary DNA strand [39]. Cleavage of each DNA strand occurs by the RuvC and HNH nuclease domains, respectively.

### 3. Designing sgRNA to Target a Genomic Target Site

For experimental purposes employing the use of Sp Cas9, crRNA and tracrRNA can be combined into a single guide RNA (sgRNA), where crRNA contains a 20 bp complementary sequence to a targeted genomic DNA region and the tracrRNA acts as a scaffold. The 20 bp complementary sequence should be followed by an “NGG” trinucleotide PAM for Sp Cas9 binding and cleavage to occur, but non-canonical PAMs (“NAG”, “NTG”, and “NGA”) are also sometimes recognized [40]. Furthermore, Cas9 may be able to bind to a sequence with as many as 10 mismatches in the spacer of crRNA [41], yet only cleaves and induces mutations at sequences with 3–5 mismatches [42]. Previous reports have termed the PAM and PAM adjacent sequence (3' half of the spacer) the seed sequence because this DNA region is scanned by the Cas9-sgRNA complex from the PAM side [43]. The seed sequence is critical for cleavage specificity and is less permissive for mismatches, while some mismatches in the non-seed region (5' half of the sgRNA target sequence) may be tolerated for Cas9 binding and DNA cleavage. Due to this nonspecific nature, selecting a target sequence that has minimal overlap with other sequences located next to a canonical and non-canonical PAM is crucial for the initial sequence recognition and specificity. Therefore, CRISPR sgRNA should be designed to target a highly unique genomic region.

### 4. Tools to Design Highly Specific sgRNA

To design specific sgRNA, it is critical to evaluate the number of potential off-target sequences in the genome. There are now a growing number of web tools available to aid in the identification of targetable sites for CRISPR sgRNAs and evaluate their specificity (Table 1). Each sgRNA design tool has its own unique properties and algorithms, but most first search for PAM sequences in a specified genomic sequence or in a region defined by an accession number according to the target region size. CRISPR Design [6,44], selects sgRNA from an input sequence up to 250 bp, while CRISPRdirect [45,46], ZiFit Targeter [47,48], Cas9 design [49,50], E-CRISP [51,52], and CHOPCHOP [53,54] can perform searches from much longer target regions.

**Table 1.** Software to design CRISPR (clustered regularly interspersed short palindromic regions) sgRNA (single guide RNA).

Tool	Organism	Input (Length)	off-Target Sites	Reference
ZiFiT	Hs, Rn, Mm, Dr, Dm, Ce, Aa, Ec	Target sequence (<1000 bp)	Mismatches	[48]
CRISPR Design	>15 species	Target sequence (<250 bp)	Mismatches	[6]
Cas9 design	>10 species	Target sequence (>10 kbp)	Mismatches	[50]
E-CRISP	>15 species	Target sequence (>10 kbp), gene name	Mismatches	[52]
CasOT	Any species	Target sequence (>10 kbp)	Mismatches	[55]
Cas-OFFinder	>10 species	Designed sgRNA (10–25 nt)	Mismatches, insertion and deletions	[56]
CHOPCHOP	>20 species	Target sequence (>10 kbp)	Mismatches	[54]
GT-Scan	>20 species	Target sequence (<4000 bp)	Mismatches	[57]
sgRNAscas9	Any species	Target sequence (>10 kbp)	Mismatches	[58]
CRISPR-P	>20 plants	Target sequence (<5 kbp)	Mismatches	[59]
COSMID	Hs, Mm, Rn, Ce, Mam, Dr	Designed sgRNA (10–55 nt)	Mismatches, insertions and deletions	[60]
sgRNA Designer	Hs, Mm	Target sequence (<10 kbp), gene ID	N.A.	[61]
iGEATs	Hs, Mm	Chromosomal locus, target sequence (<25 kb), gene name, gRNA	N.A.	[17]
CRISPRdirect	>15 species	Target sequence (<10 kbp)	Mismatches, insertions and deletions	[46]
CRISPR-ERA	Hs, Mm, Rn, Dr, Dm, Ce, Sc, Ec, Bs	Target sequence (<5 kbp), gene name	Mismatches	[62]
Protospacer Workbench	Any species	Target sequence (>10 kbp), gene name	Mismatches	[63]

Hs: *Homo sapiens*, Rn: *Rattus norvegicus*, Mm: *Mus musculus*, Dr: *Danio rerio*, Dm: *Drosophila melanogaster*, Aa: *Aedes aegypti*, Ec: *Escherichia coli*, Sc: *Saccharomyces cerevisiae*, Bs: *Bacillus subtilis*, Ce: *Caenorhabditis elegans*, Mam: *Macaca mulatta*.

The web tool CRISPRdirect searches not only the entire sgRNA sequence for potential off-target sites, but also the seed sequences, such as the first 8 and 12 nt in addition to PAM. This way, users can choose their preferred sgRNA with the minimal off-target sites in the seed region. Perl based stand-alone software CasOT [64] and sgRNAscas9 [65] also consider the number and position of mismatches, for instance, whether the position of the mismatches are located in the non-seed or seed region for ranking the potential off-target sites. In addition, previous studies have shown that sgRNA sequences with extremely high or low GC content (*i.e.*, >80% or <20%) are less effective against their targets [66]. To avoid sgRNAs with skewed GC content, CRISPRdirect provides the GC content of the designed sgRNAs.

Further analysis of potential off-target sites for candidate sgRNAs can be searched using web tools, such as Cas-OFFinder [67] and COSMID [68]. These web tools search for 20 nt target sequences

adjacent to PAMs (e.g., “NGG”, “NRG”) from an input sequence and then output the number of off-target sites and sequences.

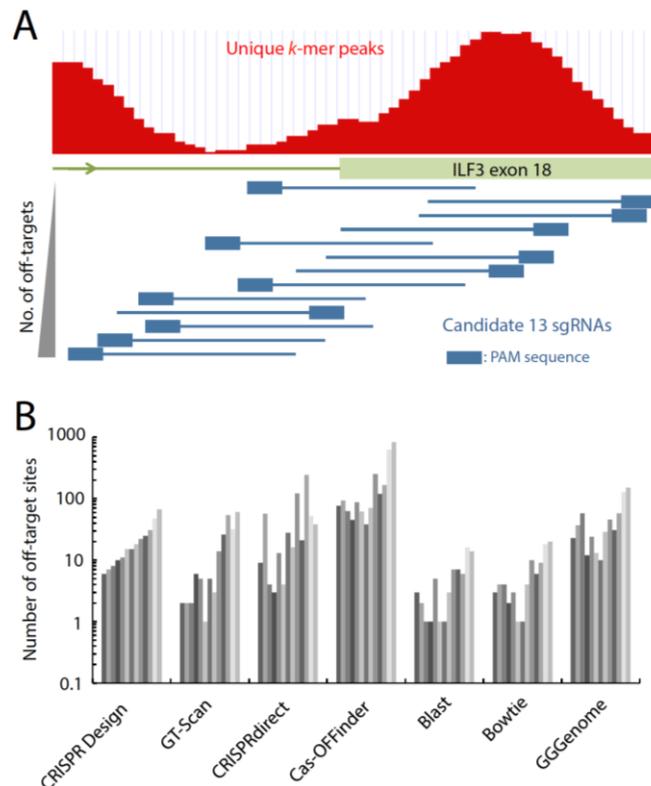
Off-target sites are normally searched based on mismatches rather than indels, as shown in Table 1; however, Bao *et al.* showed that sgRNA with a few indels may induce cleavage [69]. Therefore, it is recommended to search for potential off-target sites with indel allowances in addition to base-pair mismatches. CRISPRdirect, Cas-OFFinder, and COSMID consider both indels and mismatches for searching off-target sites and show the positions of the mismatches and gaps, which may help to predict the potency of off-target mutagenesis. In addition to CRISPR sgRNA dedicated software, sequence search tools (such as Blast, Blat, GGenome, and TagScan) or short read mapping software (such as Bowtie and BWA) might be useful to search for potential off-target sites. Although these general sequence aligners do not consider the PAM or the seed when scoring similar sequences to sgRNA, they are easy to use for quickly checking the degree to sequence specificity.

Most of the online tools have a low throughput to design multiple sgRNAs, and available organisms are limited. However, recently developed off-line software, such as CasOT [55], sgRNAs9 [58], and Protospacer Workbench [63], have greater flexibility to select targets from any organism where a genomic sequence is available. Offline software provide greater flexibility and computational speed, but some, such as CasOT [55] or sgRNAs9 [58], require programming knowledge and familiarity with a command line interface.

On-target DNA cleavage activity of sgRNA is also critical when it comes to designing sgRNAs. Apart from the GC content of sgRNAs mentioned above, there is a preference for the nucleotide immediately adjacent to PAM to be “G”, and not “C” [61,66,70,71]. The epigenetic state of the target genome was indicated to have an influence on Cas9 accessibility and subsequent cleavage activity [71], however, CpG methylation was reported separately to not significantly alter this activity [6].

In order to simulate the process of selecting sgRNAs, we designed sgRNAs targeting the exon 18 of the human dsRNA binding protein ILF3 gene on chromosome 19 for a hypothetical gene knockout experiment. First, we used iGEATs (*in silico* Genome Editing and Analysis Tools) [72] to visually identify a unique and targetable region across the exon 18 of ILF3 gene, based on the peak height of short *k*-mer sequences that mapped to the human genome (Figure 2A). We identified that the beginning of exon 18 would serve as a good target region because unique *k*-mer peaks are abundant. Then, we used the CRISPR Design tool to extract candidate sgRNA sequences from the query target sequencing (around 60 bp region) around the splicing acceptor site of exon 18 of ILF3. We selected all 13 sgRNA sequences from the CRISPR Design output, and calculated the number of potential off-target sites by GT-Scan, CRISPRdirect, and Cas-OFFinder. We also used general sequence search tools, blastn+ [73], Bowtie [74], and GGenome to search for the number of potential off-target sites, including insertions or deletions (Figure 2B). Although the computational speed to obtain results was relatively slow, the CRISPR design tool showed not only the number of off-target sites, but also those within gene coding regions. GT-Scan and CRISPRdirect could rapidly calculate the number of the off-target sites, but could not distinguish whether potential off-target sites exist within gene regions. Notably, different sgRNA design tools identify the same 13 sgRNA sequences within the given query sequence, and the ranking of the sgRNAs were relatively conserved among the four sgRNA design tools we tested. The total number of the identified off-targets greatly varied among different tools, presumably due to the different algorithms and/or parameter stringencies (Figure 2B). Furthermore, the

trends of the number of potential off-targets were also conserved among sequence search tools, blastn+ (command line mode, “blastn-short” task,  $e$ -value  $< 1.0$ ), Bowtie (command line mode, 3 bp mismatch allowance, seed length 16 bp), and GGGenome (3 bp mismatch allowance). Since blastn+ and Bowtie provided less potential off-target sites compared with the sgRNA design tools, some potential off-target sites may have been excluded. These results suggest that the full identification of all off-target sites is challenging by using some of general sequencing mapping software. To determine which off-target prediction method is the most reliable, a more comprehensive comparison of the various tools along side-by-side with actual experimental data is needed.



**Figure 2.** Tendency of identified off-target sites by sgRNA design and sequence search tools. **(A)** To demonstrate the sgRNA design process, we selected the human *ILF3* gene as a target. To identify uniquely targetable region of sgRNAs, genomic region around the splicing acceptor site of exon 18 and peaks of unique  $k$ -mer sequences are shown as red histogram. Thirteen candidate sgRNA sequences (blue arrow with a box, indicating the position of PAM sequence) were identified within the region and the identified sgRNAs are ordered based on the rank provided by the CRISPR design tool; **(B)** With the same set of 13 sgRNAs as shown in **(A)**, we calculated the number of potential off-target sites in human genome (hg19) by using various tools as indicated. For the calculation, the following options are used: GT-Scan, allowing three mismatches; CRISPRdirect, potential targets with 12 nt plus PAM; Cas-OFFinder, allowing three mismatches, two DNA bulge size, and two RNA bulge size; stand alone Blastn+, “blastn-short” task,  $e$ -value  $< 1.0$ ; stand alone Bowtie, allowing three mismatches, seed length 16 nt; GGGenome, allowing three mismatches.

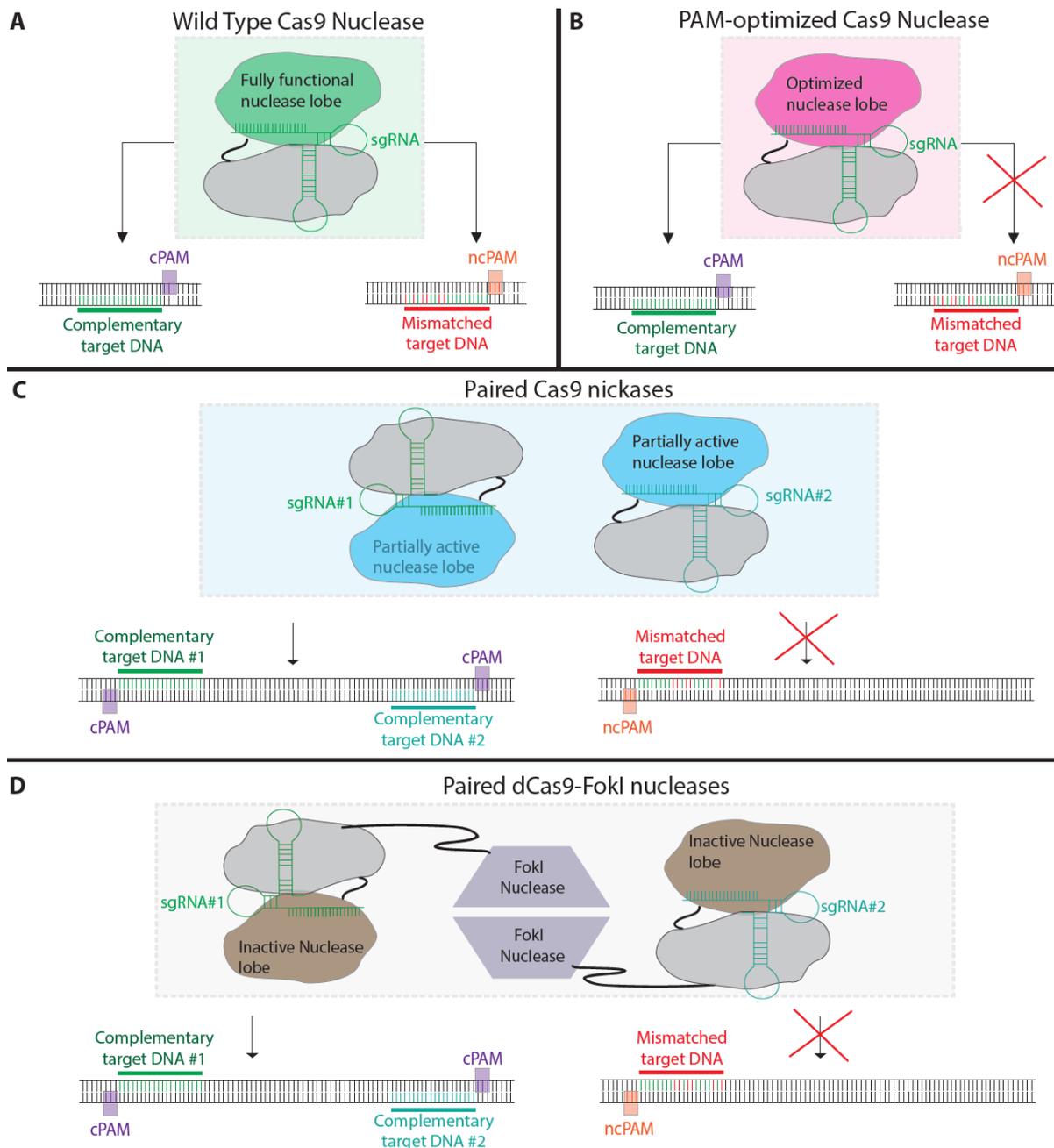
## 5. Modified Nucleases to Reduce off-Target Mutagenesis

Modulating Cas9 activity through the substitution of amino acid residues involved in its enzymatic function could fine tune substrate recognition and reduce nonspecific DNA cleavage. Alternatively, by completely abolishing the enzymatic activity of Cas9 while retaining its ability to target DNA, a higher fidelity endonuclease could be fused to a deactivated Cas9 homing module to increase on-target cleavage.

In one strategy, the Cas9 PAM recognition domain can be modified to reduce the number of potential off-target mutagenesis. An *in vitro* assessment of recombinant Cas9 by single molecule imaging revealed that Cas9 complexed with sgRNA scans complementary DNA sequences depending on the presence of possible PAM recognition sites [43]. Sp Cas9 typically recognizes “NGG” but can also use non-canonical PAMs, thereby increasing off target cleavage (Figure 3A). In a recent report, by Joung and colleagues, Sp Cas9 was engineered with a D1135E mutation that increased the specificity of on-target cleavage by reducing non-canonical PAM recognition (Figure 3B) [75]. In addition, the group modified Sp Cas9 to recognize different PAM sequences with different lengths. Such engineering may be able to fine tune Cas9 activity and greatly enhance specificity, which is defined as the ratio of on-site cleavage to off-site cleavage. More recently, Davis *et al.* [76] showed that the suppression of off-target cleavage by insertion of an engineered intein (which is fused with estrogen receptor ligand-binding domain) into Sp Cas9. The addition of 4-hydroxytamoxifen induced protein splicing to remove the engineered intein and activate Cas9. As a result, DNA cleavage specificity improved as high as 25-fold compared with wild-type Cas9.

Two additional mutagenesis approaches have been used to mitigate CRISPR Cas9 off target cleavage activity by attenuating or abolishing its nuclease function. For partial mutagenesis, a known aspartate to alanine mutation (D10A) inactivates the RuvC nuclease domain while leaving the HNH domain intact [1]. The resulting nickase activity cleaves only one strand of the dsDNA, creating a small gap as opposed to a complete blunt dsDNA break, and is subsequently fixed by a less error prone base excision repair (BER) pathway which involves multiple enzymes that recognize and repair the damaged DNA. Single nicking enhances HR events, but is much less effective than DSB mediated HR. Combining two closely spaced sgRNAs, offset by less than 100 bp, that target opposite strands of a DNA target sequence can lead to the efficient induction of DSBs and HR (Figure 3C) [6,77]. The general consensus from several groups has been that the double nickase strategy efficiently cleaves on-target sites as well as the WT Cas9, while lowering off-target mutations [77–79]. However, even this method is not completely free of risks as point mutations can result from nickase activity [80].

Fusing a catalytically dead Cas9 protein (D10A and H840A mutations) with a Fok I nuclease, similarly to what has been done with ZFNs [81,82] and TALENs, can add another level of control during genome editing. In doing so, DNA targeting by Cas9-sgRNA is retained and cleavage activity occurs only after FokI dimerization [80,83] To function, two unique sgRNAs targeting DNA sequences, which are spaced approximately 17 base pairs apart, can bring together Cas9-FokI chimeras into close enough proximity of each other to allow for dimerization and subsequent DNA cleavage (Figure 3D). Thus, it is possible to maintain the DNA homing characteristics of CRISPR Cas9 while, at the same time, replacing its off-target behavior with a more specific FokI nuclease [80]. The downside to using FokI is that it is significantly larger than Cas9, meaning that the transduction efficiency of the plasmid DNA could be compromised.



**Figure 3.** Modified Cas9 strategies to reduce off-target mutagenesis. **(A)** Wild type Sp Cas9 nuclease can bind and cleave either sgRNA-complementary (green) or -mismatched (red) DNA with canonical PAM (cPAM) or noncanonical PAM (ncPAM), respectively; **(B)** PAM-optimized Cas9 nuclease induces cleavage specifically to cPAM complementary DNA; **(C)** A pair of SpCas9 nickases (*i.e.*, D10A mutant) must bind to two closely spaced DNA sequences on opposite strands in order to induce single strand nicking on each strand and result in full cleavage of the target DNA. The binding of a Cas9 nickase to a non-specific site will induce nicking but not a double strand break; **(D)** Deactivated Sp dCas9 (*i.e.*, D10A and H840A mutant) fused with a FokI nuclease must bind to two closely spaced DNA sequences in order for dimerization of the FokI nuclease and cleavage to occur. Nonspecific binding at an off-target site will not induce cleavage.

While gene editing experiments to date have focused on Sp Cas9, *Staphylococcus aureus* (Sa) Cas9 has been reported to have high cleavage activity, comparable to that of Sp Cas9. Unlike the “NGG” PAM recognition of Sp Cas9, Sa Cas9 has a longer PAM consisting of “NNGRRN” meaning that the number of off-target candidates could be less [75,84]. However, a larger sample size is needed before any conclusions can be drawn.

## 6. Cas9 Delivery and Expression

To express Cas9 protein and sgRNA, DNA plasmid transfections have been commonly employed in various cell lines for experiments examining on- and off-target cleavage. However, the overexpression of both Cas9 protein and sgRNA by DNA plasmid transfection has been shown to induce off-target mutagenesis when higher amounts of DNA are transfected [6]. Thus, there appears to be an undefined threshold where the levels of Cas9 protein and sgRNA could become “toxic” and lead to a higher occurrence of off-target effects. Intuitively, this makes sense because increasing the number of interaction events increases the probability that Cas9 may bind and cleave an off-target site.

Other methods to temporally express Cas9 in cells include the use of an inducible CRISPR (iCRISPR) Cas9 expression system, and delivery by co-transfecting Cas9 mRNA and sgRNA or Cas9 protein pre-loaded with sgRNA. Stable iCRISPR expressing human pluripotent cells, reported by a study led by Huangfu [85], can transiently express Cas9 under the control of a tetracycline responsive promoter. By transfecting sgRNAs into these cells, efficient knockout of genes could be performed with reduced off-target cleavage. Another approach to limiting Cas9 expression by DNA transfection is to simultaneously target the Cas9 plasmid with an additional sgRNA to prevent prolonged protein expression of the nuclease [86].

Alternatively, because RNA and protein are unstable by nature and have relatively short half-lives, their direct delivery into cells may also reduce the transient time that Cas9 exists, and therefore, limit off-target effects. Indeed, Kim and colleagues [87] showed that the specificity of on-target cleavage *versus* off target cleavage is up to 13 times greater when using Cas9 protein compared with Cas9 plasmid delivery. When the authors examined Cas9 protein levels over a time course of 72 h, protein was undetectable after 24 h, whereas plasmid transfection resulted in peak Cas9 protein levels at 24 h with high levels still detected up to 72 h [87]. Altogether, these results suggest that prolonged Cas9 and sgRNA expression can lead to accumulating mutations.

Recently, chemical modifications at the 3' end of sgRNAs with 2'-thionocarbonate-protected nucleoside phosphoramidites have been reported to enhance cleavage activity and improve the specificity of some sgRNA [88]. In addition, it is reported that short sgRNA (*i.e.*, 17 nt) have higher specificity, as it is less tolerated to mismatches to induce DNA cleavage [89]. Therefore, modifying sgRNA is another option to reduce non-specificity.

## 7. Assessing off-Target Mutagenesis

A robust evaluation of potential off-target sites is necessary to fully assess the extent of non-specific cleavage by Cas9 [90,91]. There are three main categories of assessing methods: (i) search and assess off-target mutagenesis at locations similar to the on-target sgRNA sequence; (ii) assess the genomic regions where DSBs have been induced; and (iii) search the whole genome to check for genomic

integrity. In each category, several strategies are available, which vary in sensitivity and coverage of both biased and unbiased sequences. Furthermore, whether to analyze bulk cultured cells or cells isolated by subcloning should be considered to determine the types of mutations that can arise in individual cells types.

For the first category, potential off-target sites are predicted using appropriate software as mentioned above, and then assessed by a mutation detection assay, such as T7EI (T7 endonuclease I) assay or deep sequencing of the PCR products (also known as amplicon seq). For the T7EI assay, PCR primers to amplify the predicted cleavage sites from genomic DNA are designed to be approximately 500 to 1000 bp apart. The amplicons are denatured and slowly re-annealed so that strands containing indels will anneal to uncleaved strands creating mismatches that the T7E1 endonuclease will cleave (similarly in principle, the CEL-I assay or SURVEYOR assay can be used). Any cleavage products run on an electrophoresis are positive indicators of Cas9 activity. Many laboratories use this technique as it is inexpensive and no special equipment is required, but its sensitivity is low and coverage is limited due to specific primers needed for each predicted target sequence (*i.e.*, deletions larger than primer binding sites cannot be detected). In addition, homozygous mutations can not be detected, unless non-mutated PCR products are spiked-in. Sanger sequencing of PCR amplicons makes it possible to detect indels with higher sensitivity but suffers from limited coverage of the chosen primer sets.

To look at the off-target binding of Cas9 in an exogenously established system, Church and colleagues fused a catalytically dead Cas9 protein (dCas9) with a transcriptional activator, VP64, to screen a library of sequences that, when bound to dCas9-VP64, activate a fluorescence reporter in HEK293T cells [78]. This study demonstrated the off-target binding affinity of Cas9 protein with sgRNAs containing up to three mismatches. A similar study led by Sander *et al.* [42] used an EGFP disruption assay to test the specificity of CRISPR Cas9 cleavage with sgRNAs harboring up to five mismatches at different positions in the sgRNA. The authors used HEK293T cells stably expressing EGFP, and examined gRNA candidates that diminished EGFP expression.

SELEX (systematic evolution of ligands by exponential enrichment) is an *in vitro* method that investigates the binding of Cas9, or any nuclease, to a library of oligonucleotides to anticipate potential off-target sites under controlled conditions [41,92]. Genomic DNA sequences that match those of the bound oligonucleotides can then be searched.

An interaction approach by chromatin immunoprecipitation sequencing (ChIP-seq) takes advantage of binding by dCas9 to DNA sequences for assessing potential off target sites [41,93,94]. However, it is worth mentioning that a comparison between the off-target identification using CHiP Seq with catalytically dead Cas9 and catalytically active Cas9 did not correlate well, meaning that the binding affinity of Cas9 to target sites does not always lead to cleavage [93].

As the second category of off-target analysis, a DSB event itself can be tagged or captured to look for potential off-target mutagenesis by deep sequencing. One such method is to use integration-defective lentiviral vectors (IDLVs) that are incorporated into DSBs induced by Cas9 [95]. The vector generally contains a selection marker such as puromycin resistance, which can be used to enrich the population of cells with incorporated IDLVs. The technique has been used for ZFNs, TALENs, and CRISPR Cas9 but has a sensitivity of 0.5% and may be underestimating the total number of DSBs.

Similarly to map genome-wide DSBs, BLESS (direct *in situ* breaks labeling, enrichment on streptavidin and next-generation sequencing) captures a snapshot of DSBs in fixed cells by labeling the DSB sites with biotinylated linkers and deep-sequencing after PCR amplification of the captured DNA

fragments [96]. The technique has been successfully applied to assessing Sa Cas9 off-target cleavage in 293FT cells, which are derived from HEK293 cells [84].

In a more sensitive assay to capture DSBs in live cells, GUIDE-seq (Genome-wide, unbiased identification of DSBs enabled by sequencing) is a method that tags DSBs by incorporating short phosphorylated double-stranded oligodeoxynucleotides into them [97]. The regions are amplified by PCR and then read by next generation sequencing. This technique is sensitive enough to detect mutagenesis that occurs at a rate of 0.1%. To detect DSB mediated chromosomal translocation, the HTGTS (high-throughput, genome-wide, translocation sequencing) method has been reported [94]. Linear amplification mediated (LAM)-PCR amplifies the region near the nuclease target site from one direction, and searches for translocated chimeric sequencing reads from deep sequencing data. Another unbiased approach called digenome-seq (digested genome) subjects genomic DNA, extracted from untreated or Cas9 transfected cells, to cleavage by recombinant Cas9 and sgRNA *in vitro* [98]. By comparing both sample reads from whole genome sequencing data, a pattern will emerge where uncleaved and cleaved sites can be distinguished. Like GUIDE-seq, the sensitivity of the assay is as high as 0.1%, however, high-coverage WGS data is required to fully achieve the sensitivity. The overall genomic stability in edited cells can be affected by both sgRNA-dependent and -independent events. For example, sgRNA-independent abnormalities can be caused from various forms of stress induced by experimental conditions such as transduction of Cas9-sgRNA complexes.

As the third category of the assay to detect off-target mutagenesis, genomic integrity can be examined by various methods. This category of the assays is much more relevant for clinical applications, as proof of genomic integrity can be a minimal requirement. Genomic abnormalities induced by CRISPR sgRNAs and/or other means can be highly variable in terms of their sizes. For example, chromosomal abnormalities, such as deletions, duplications, inversions, or translocations of larger than 1 Mb, karyotyping is a sensitive and established assay to detect minor cell populations with chromosomal abnormalities (*i.e.*, >10%–20% of cells). However, the assay relies on imaging of chromosomes, and resolution is very low.

The cleavage of DNA at two separate sites in the genome may lead to a loss of large DNA segments, as long as 1 Mbp [99]. For this kind of large deletion, the DNA copy number variance (CNV) should be examined in both the parental and CRISPR Cas9 treated cells [17]. Since CNV analysis uses array technology to detect genomic copy number in a cell population, resolution is higher than karyotyping (~1 kb), but sensitivity is low so that at least >30%~40% of cell population must have a similar CNV to be detected reproducibly.

Investigation of mutations at single base-pair resolution requires the use of deep sequencing. Exome sequencing is a method to capture protein coding regions of the genome and analysis by deep sequencing. Coverage of regions is limited, however, higher coverage can be achieved than whole genome sequencing at a lower cost. In addition, a consequence of the detected mutation can be inferred more easily than other methods, as effect in amino acid coding can be evaluated. Several groups have been using exome sequencing to identify potential off-target mutagenesis in genome edited cells [17].

For a more comprehensive survey of off-target mutagenesis, whole genome sequencing (WGS) has been performed for cell lines [100–102] and mice [103] edited by TALENs and CRISPR Cas9. So far, no severe off-target mutagenesis attributed to the nucleases has been found. WGS is a powerful tool to detect genome-wide single nucleotide variations and small indels, but to detect larger deletions

(*i.e.*, CNV), specific analysis pipelines or algorithms need to be employed. This may have led to the exclusion of real target sites that were confused with artifacts during sequencing analysis or a lack of sensitivity to detect low frequency mutants that occur in less than 10% of clones. In sum, the combination of several off-target assays should be considered to fully assess the risk of off-target mutagenesis.

## 8. Off-Target Analysis Based on Genome Editing Objectives

How stringently off-target cleavage should be considered when using genome editing tools will depend on the research aim of the investigating scientist. For basic research (*i.e.*, gene knockout, reporter knock-in) using cell lines, off-target mutagenesis should be minimized by careful design of sgRNAs, but it is not an absolute requirement so long as the scientific conclusions are supported by multiple sgRNAs or clones, as it is unlikely that the same off-target mutagenesis will occur in multiple clones. For the clinical setting, using *ex vivo* gene therapy approaches, sgRNA must be carefully designed to have minimum off-target cleavage risks. However, even with an optimized experimental design, initial clinical trials should rigorously examine the genomic integrity of the edited cells. Detection of DSB sites by massive parallel sequencers is a powerful approach, however, the genomic integrity cannot be guaranteed whenever there is a DSB involved. Classical karyotyping, CNV analysis, and exome sequencing, or whole genome sequencing might be required to show that the genomic DNA in the genome edited cells is intact. For an *in vivo* gene therapy approach, a complete and rigorous design of sgRNA is required. However, verifying the genomic integrity of the treated cells *in vivo* is much more challenging.

## 9. Target Cells Matter

There have been conflicting reports in the literature about the detection of off-target cleavage by engineered nucleases, in particular TALENS and CRISPR Cas9, in different cell types (Table 2). Interestingly, in the majority of studies that detected off-target mutagenesis, nearly all of the experiments were performed in cancer cell lines such as HEK293, U2OS, and K562. On the other hand, in reports where off-target mutations were undetected, pluripotent stem cells (ES and iPS cells) were used. It is important to note that while these studies cannot be compared side by side because the experimental parameters and methods used to detect the off-target mutations were not identical, the different results may indicate a higher propensity of observed off-target mutations to occur in cancer cell lines where the DNA repair machinery has been compromised. In fact, Joung and colleagues identified DSB “hot spots”, or regions that are likely to have DSBs, independent of CRISPR Cas9 in HEK293 and U2OS cells [97]. On the contrary, several lines of studies identified pluripotent stem cell specific “genomic guardians”, such as *Zscan4* [104], to maintain genomic stability. Future studies will elucidate whether there is a correlation between off-target mutagenesis and the cell type being edited.

**Table 2.** Summary of studies detecting off-target mutagenesis using different cell lines.

Cell Types	Target Gene	Programmable Nuclease Used	Detection Assay of off-Target Mutagenesis	off-Target Mutagenesis Detected	Ref.
U2OS, HEK293 and K562 cells	<i>VEGFA, EMX1, RNF2, FANCF</i>	CRISPR	EGFP reporter, T7EI	Yes	[42]
293T	<i>EMX1</i>	TALENs, CRISPR	amplicon seq, Surveyor assay	Yes	[6]
293T	<i>HBB, CCR5</i>	CRISPR	T7EI, Sanger sequencing	Yes	[105]
293T	<i>HBB, CCR5</i>	CRISPR	T7EI, Sanger sequencing	Yes	[69]
293 and U2OS cells	<i>VEGFA, EMX1, FANCF, RNF2</i>	CRISPR	GUIDE-seq, amplicon seq (AMP-based seq)	Yes	[97]
HAP1 cells, K562 cells,	<i>HBB, VEGFA</i>	CRISPR	Digenome-seq	Yes	[98]
293T cells, A549 cells	<i>RAG1, C-MYC, ATM</i>	TALENs, CRISPR	HTGTS method	Yes	[94]
iPS cells	<i>A1AT</i>	ZFNs	CGH, SNP array, exome seq	No	[106]
myoblasts	<i>DMD</i>	TALENs	exome seq	No	[107]
iPS cells	<i>PPP1R12C</i>	TALENs, CRISPR	whole genome seq	No	[100]
ES cells, iPS cells	<i>SORT1, LINC00116</i>	TALENs, CRISPR	whole genome seq	No	[102]
iPS cells	<i>HBB</i>	TALENs, HDAdV mediated HR	whole genome seq	No	[101]
293FT cells, iPS cells	<i>PPP1R12C, AKT2, CDK19, ATP6AP2, SLC35A2</i>	CRISPR	amplicon seq	No	[108]
iPS cells	<i>DMD</i>	TALENs and CRISPR	T7EI, amplicon seq, Karyotyping, CNV analysis, exome seq	No	[17]
iPS cells	<i>PLN</i>	TALEN	exome seq	No	[109]
iPS cells	<i>F8</i>	CRISPR	amplicon seq	No	[21]

## 10. Conclusions

In conclusion, there are several considerations to take into account in order to minimize off-target mutagenesis when using the CRISPR Cas9 system. The sgRNA design is critical to avoid similar sequences, and numerous online tools are available to select highly specific sgRNA complementary sequences. A modified Cas9 nuclease can also be selected for targeting purposes, such as the paired D10A Cas9 approach over the wild type nuclease. In addition, further enhancements to the PAM recognition motif may help to increase canonical PAM specificity while reducing the non-canonical PAM promiscuity. Finally, the detection method used will be important to discover low frequency

mutagenesis at off-target sites. However, in considering these options, the need for specificity will depend on the research aims. Basic researchers using cancer cell lines will have less interest in conducting whole genome sequencing in order to identify off-target mutagenesis, whereas a researcher aiming to correct a genetic disease for clinical applications may be required to minimize off-target risks, such as tumorigenesis, in treated patients.

### Acknowledgments

The authors apologize to all whose work could not be cited due to space limitations. Akitsu Hotta is supported in part by JSPS KAKENHI grant (15H05581) and AMED Research Center Network for Realization of Regenerative Medicine grants.

### Author Contributions

Kentaro Ishida, Peter Gee, and Akitsu Hotta jointly wrote the manuscript.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821.
2. Bhaya, D.; Davison, M.; Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Ann. Rev. Genet.* **2011**, *45*, 273–297.
3. Wiedenheft, B.; Sternberg, S.H.; Doudna, J.A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **2012**, *482*, 331–338.
4. Cho, S.W.; Kim, S.; Kim, J.M.; Kim, J.-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **2013**, *31*, 230–232.
5. Mali, P.; Yang, L.; Esvelt, K.M.; Aach, J.; Guell, M.; DiCarlo, J.E.; Norville, J.E.; Church, G.M. RNA-guided human genome engineering via Cas9. *Science* **2013**, *339*, 823–826.
6. Hsu, P.D.; Scott, D.A.; Weinstein, J.A.; Ran, F.A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E.J.; Wu, X.; Shalem, O.; *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **2013**, *31*, 827–832.
7. Hotta, A.; Yamanaka, S. From Genomics to Gene Therapy: iPS Cells Meet Genome Editing. *Annu. Rev. Genet.* **2015**, *49*, doi:10.1146/annurev-genet-112414-054926.
8. Tzur, Y.B.; Friedland, A.E.; Nadarajan, S.; Church, G.M.; Calarco, J.A.; Colaiácovo, M.P. Heritable custom genomic modifications in *Caenorhabditis elegans* via a CRISPR-Cas9 system. *Genetics* **2013**, *195*, 1181–1185.
9. Friedland, A.E.; Tzur, Y.B.; Esvelt, K.M.; Colaiácovo, M.P.; Church, G.M.; Calarco, J.A. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* **2013**, *10*, 741–743.
10. Chang, N.; Sun, C.; Gao, L.; Zhu, D.; Xu, X.; Zhu, X.; Xiong, J.-W.; Xi, J.J. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **2013**, *23*, 465–472.

11. Platt, R.J.; Chen, S.; Zhou, Y.; Yim, M.J.; Swiech, L.; Kempton, H.R.; Dahlman, J.E.; Parnas, O.; Eisenhaure, T.M.; Jovanovic, M.; *et al.* CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **2014**, *159*, 440–455.
12. Wang, H.; Yang, H.; Shivalila, C.S.; Dawlaty, M.M.; Cheng, A.W.; Zhang, F.; Jaenisch, R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **2013**, *153*, 910–918.
13. Li, D.; Qiu, Z.; Shao, Y.; Chen, Y.; Guan, Y.; Liu, M.; Li, Y.; Gao, N.; Wang, L.; Lu, X.; *et al.* Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 681–683.
14. 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*, 684–686.
15. Yang, D.; Xu, J.; Zhu, T.; Fan, J.; Lai, L.; Zhang, J.; Chen, Y.E. Effective gene targeting in rabbits using RNA-guided Cas9 nucleases. *J. Mol. Cell Biol.* **2014**, *6*, 97–99.
16. Niu, Y.; Shen, B.; Cui, Y.; Chen, Y.; Wang, J.; Wang, L.; Kang, Y.; Zhao, X.; Si, W.; Li, W.; *et al.* Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* **2014**, *156*, 836–843.
17. Li, H.L.; Fujimoto, N.; Sasakawa, N.; Shirai, S.; Ohkame, T.; Sakuma, T.; Tanaka, M.; Amano, N.; Watanabe, A.; Sakurai, H.; *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*, 143–154.
18. Li, H.L.; Nakano, T.; Hotta, A. Genetic correction using engineered nucleases for gene therapy applications. *Dev. Growth Differ.* **2014**, *56*, 63–77.
19. Xie, F.; Ye, L.; Chang, J.C.; Beyer, A.I.; Wang, J.; Muench, M.O.; Kan, Y.W. Seamless gene correction of  $\beta$ -thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and *piggyBac*. *Genome Res.* **2014**, *24*, 1526–1533.
20. Schwank, G.; Koo, B.-K.; Sasselli, V.; Dekkers, J.F.; Heo, I.; Demircan, T.; Sasaki, N.; Boymans, S.; Cuppen, E.; van der Ent, C.K.; *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **2013**, *13*, 653–658.
21. Park, C.-Y.; Kim, D.H.; Son, J.S.; Sung, J.J.; Lee, J.; Bae, S.; Kim, J.-H.; Kim, D.-W.; Kim, J.-S. Functional correction of large factor VIII gene chromosomal inversions in hemophilia A patient-derived iPSCs using CRISPR-Cas9. *Cell Stem Cell* **2015**, *17*, 213–220.
22. 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*, 659–662.
23. Yin, H.; Xue, W.; Chen, S.; Bogorad, R.L.; Benedetti, E.; Grompe, M.; Kotliansky, V.; Sharp, P.A.; Jacks, T.; Anderson, D.G. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* **2014**, *32*, 551–553.
24. Long, C.; McAnally, J.R.; Shelton, J.M.; Mireault, A.A.; Bassel-Duby, R.; Olson, E.N. Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* **2014**, *345*, 1184–1188.
25. Ramanan, V.; Shlomai, A.; Cox, D.B.T.; Schwartz, R.E.; Michailidis, E.; Bhatta, A.; Scott, D.A.; Zhang, F.; Rice, C.M.; Bhatia, S.N. CRISPR/Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus. *Sci. Rep.* **2015**, *5*, doi:10.1038/srep10833.

26. Seeger, C.; Sohn, J.A. Targeting hepatitis B virus with CRISPR/Cas9. *Mol. Ther. Nucleic Acids* **2014**, *3*, doi:10.1038/mtna.2014.68.
27. Ebina, H.; Misawa, N.; Kanemura, Y.; Koyanagi, Y. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.* **2013**, *3*, doi:10.1038/srep02510.
28. Ye, L.; Wang, J.; Beyer, A.I.; Teque, F.; Cradick, T.J.; Qi, Z.; Chang, J.C.; Bao, G.; Muench, M.O.; Yu, J.; *et al.* Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5 $\Delta$ 32 mutation confers resistance to HIV infection. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9591–9596.
29. Liao, H.-K.; Gu, Y.; Diaz, A.; Marlett, J.; Takahashi, Y.; Li, M.; Suzuki, K.; Xu, R.; Hishida, T.; Chang, C.-J.; *et al.* Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat. Commun.* **2015**, *6*, doi:10.1038/ncomms7413.
30. Cornu, T.I.; Thibodeau-Beganny, S.; Guhl, E.; Alwin, S.; Eichinger, M.; Joung, J.K.; Joung, J.K.; Cathomen, T. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol. Ther.* **2008**, *16*, 352–358.
31. Porteus, M.H.; Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **2003**, *300*, 763–763.
32. Guilinger, J.P.; Pattanayak, V.; Reyon, D.; Tsai, S.Q.; Sander, J.D.; Joung, J.K.; Liu, D.R. Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat. Methods* **2014**, *11*, 429–435.
33. Li, T.; Huang, S.; Zhao, X.; Wright, D.A.; Carpenter, S.; Spalding, M.H.; Weeks, D.P.; Yang, B. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res.* **2011**, *39*, 6315–6325.
34. Mussolino, C.; Morbitzer, R.; Lütge, F.; Dannemann, N.; Lahaye, T.; Cathomen, T. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* **2011**, *39*, 9283–9293.
35. Pattanayak, V.; Guilinger, J.P.; Liu, D.R. Determining the specificities of TALENs, Cas9, and other genome-editing enzymes. *Methods Enzymol.* **2014**, *546*, 47–78.
36. Jinek, M.; Jiang, F.; Taylor, D.W.; Sternberg, S.H.; Kaya, E.; Ma, E.; Anders, C.; Hauer, M.; Zhou, K.; Lin, S.; *et al.* Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* **2014**, *343*, doi:10.1126/science.1247997.
37. Nishimasu, H.; Ran, F.A.; Hsu, P.D.; Konermann, S.; Shehata, S.I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **2014**, *156*, 935–949.
38. Jiang, F.; Zhou, K.; Ma, L.; Gressel, S.; Doudna, J.A. STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* **2015**, *348*, 1477–1481.
39. Anders, C.; Niewoehner, O.; Duerst, A.; Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **2014**, *513*, 569–573.
40. Zhang, Y.; Ge, X.; Yang, F.; Zhang, L.; Zheng, J.; Tan, X.; Jin, Z.-B.; Qu, J.; Gu, F. Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells. *Sci. Rep.* **2014**, *4*, doi:10.1038/srep05405.
41. 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*, 677–683.

42. Fu, Y.; Foden, J.A.; Khayter, C.; Maeder, M.L.; Reyon, D.; Joung, J.K.; Sander, J.D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **2013**, *31*, 822–826.
43. Sternberg, S.H.; Redding, S.; Jinek, M.; Greene, E.C.; Doudna, J.A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **2014**, *507*, 62–67.
44. CRISPR Design. Available online: <http://crispr.mit.edu> (accessed on 14 October 2015).
45. CRISPRdirect. Available online: <http://crispr.dbcls.jp/> (accessed on 14 October 2015).
46. Naito, Y.; Hino, K.; Bono, H.; Ui-Tei, K. CRISPRdirect: Software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* **2015**, *31*, 1120–1123.
47. ZiFit. Available online: <http://zifit.partners.org/ZiFiT/> (accessed on 14 October 2015).
48. Sander, J.D.; Maeder, M.L.; Reyon, D.; Voytas, D.F.; Joung, J.K.; Dobbs, D. ZiFiT (Zinc Finger Targeter): An updated zinc finger engineering tool. *Nucleic Acids Res.* **2010**, *38*, W462–W468.
49. Cas9 Design. Available online: <http://cas9.cbi.pku.edu.cn> (accessed on 14 October 2015).
50. Ma, M.; Ye, A.Y.; Zheng, W.; Kong, L. A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes. *BioMed. Res. Int.* **2013**, *2013*, 1–4.
51. E-CRISP. Available online: <http://www.e-crisp.org/E-CRISP/> (accessed on 14 October 2015).
52. Heigwer, F.; Kerr, G.; Boutros, M. E-CRISP: Fast CRISPR target site identification. *Nat. Methods* **2014**, *11*, 122–123.
53. CHOPCHOP. Available online: <https://chopchop.rc.fas.harvard.edu> (accessed on 14 October 2015).
54. Montague, T.G.; Cruz, J.M.; Gagnon, J.A.; Church, G.M.; Valen, E. CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* **2014**, *42*, W401–W407.
55. Xiao, A.; Cheng, Z.; Kong, L.; Zhu, Z.; Lin, S.; Gao, G.; Zhang, B. CasOT: A genome-wide Cas9/gRNA off-target searching tool. *Bioinformatics* **2014**, *30*, 1180–1182.
56. Bae, S.; Park, J.; Kim, J.-S. Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **2014**, *30*, 1473–1475.
57. O’Brien, A.; Bailey, T.L. GT-Scan: Identifying unique genomic targets. *Bioinformatics* **2014**, *30*, 2673–2675.
58. Xie, S.; Shen, B.; Zhang, C.; Huang, X.; Zhang, Y. sgRNACas9: A software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. *PLoS ONE* **2014**, *9*, e100448–e100449.
59. Lei, Y.; Lu, L.; Liu, H.-Y.; Li, Sen; Xing, F.; Chen, L.-L. CRISPR-P: A web tool for synthetic single-guide RNA design of CRISPR-system in plants. *Mol. Plant* **2014**, *7*, 1494–1496.
60. Cradick, T.J.; Qiu, P.; Lee, C.M.; Fine, E.J.; Bao, G. COSMID: A web-based tool for identifying and validating CRISPR/Cas off-target sites. *Mol. Ther. Nucleic Acids* **2014**, *3*, e214–10.
61. Doench, J.G.; Hartenian, E.; Graham, D.B.; Tothova, Z.; Hegde, M.; Smith, I.; Sullender, M.; Ebert, B.L.; Xavier, R.J.; Root, D.E. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **2014**, *32*, 1262–1267.
62. Liu, H.; Wei, Z.; Dominguez, A.; Li, Y.; Wang, X.; Qi, L.S. CRISPR-ERA: A comprehensive design tool for CRISPR-mediated gene editing, repression and activation. *Bioinformatics* **2015**, *10*, doi:10.1093/bioinformatics/btv423.
63. MacPherson, C.R.; Scherf, A. Flexible guide-RNA design for CRISPR applications using Protospacer Workbench. *Nat. Biotechnol.* **2015**, *33*, 805–806.

64. CasOT. Available online: <http://eendb.zfgenetics.org/casot/> (accessed on 14 October 2015).
65. sgRNAsCas9. Available online: <http://www.biotoools.com> (accessed on 14 October 2015).
66. Wang, T.; Wei, J.J.; Sabatini, D.M.; Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **2014**, *343*, 80–84.
67. Cas-OFFinder. Available online: <http://www.rgenome.net/cas-offinder/> (accessed on 14 October 2015).
68. COSMID. Available online: <https://crispr.bme.gatech.edu> (accessed on 14 October 2015).
69. Lin, Y.; Cradick, T.J.; Brown, M.T.; Deshmukh, H.; Ranjan, P.; Sarode, N.; Wile, B.M.; Vertino, P.M.; Stewart, F.J.; Bao, G. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.* **2014**, *42*, 7473–7485.
70. Gagnon, J.A.; Valen, E.; Thyme, S.B.; Huang, P.; Akhmetova, L.; Akhmetova, L.; Pauli, A.; Montague, T.G.; Zimmerman, S.; Richter, C.; *et al.* Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS ONE* **2014**, *9*, e98186.
71. Chari, R.; Mali, P.; Moosburner, M.; Church, G.M. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat. Methods* **2015**, *12*, 823–826.
72. iGEATS. Available online: <https://apps.cira.kyoto-u.ac.jp/igeats/index.html> (accessed on 14 October 2015).
73. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, doi:10.1186/1471-2105-10-421.
74. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **2009**, *10*, doi:10.1186/gb-2009-10-3-r25.
75. Kleinstiver, B.P.; Prew, M.S.; Tsai, S.Q.; Topkar, V.V.; Nguyen, N.T.; Zheng, Z.; Gonzales, A.P.W.; Li, Z.; Peterson, R.T.; Yeh, J.-R.J.; *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **2015**, *523*, 481–485.
76. Davis, K.M.; Pattanayak, V.; Thompson, D.B.; Zuris, J.A.; Liu, D.R. Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat. Chem. Biol.* **2015**, *11*, 316–318.
77. Ran, F.A.; Hsu, P.D.; Lin, C.-Y.; Gootenberg, J.S.; Konermann, S.; Trevino, A.E.; Scott, D.A.; Inoue, A.; Matoba, S.; Zhang, Y.; *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **2013**, *154*, 1380–1389.
78. Mali, P.; Aach, J.; Stranges, P.B.; Esvelt, K.M.; Moosburner, M.; Kosuri, S.; Yang, L.; Church, G.M. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **2013**, *31*, 833–838.
79. Shen, B.; Zhang, W.; Zhang, J.; Zhou, J.; Wang, J.; Chen, L.; Wang, L.; Hodgkins, A.; Iyer, V.; Huang, X.; *et al.* Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat. Methods* **2014**, *11*, 399–402.
80. Tsai, S.Q.; Wyvekens, N.; Khayter, C.; Foden, J.A.; Thapar, V.; Reyon, D.; Goodwin, M.J.; Aryee, M.J.; Joung, J.K. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **2014**, *32*, 569–576.
81. Wang, J.; Friedman, G.; Doyon, Y.; Wang, N.S.; Li, C.J.; Miller, J.C.; Hua, K.L.; Yan, J.J.; Babiarz, J.E.; Gregory, P.D.; *et al.* Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res.* **2012**, *22*, 1316–1326.

82. Kim, E.; Kim, S.; Kim, D.H.; Choi, B.-S.; Choi, I.-Y.; Kim, J.-S. Precision genome engineering with programmable DNA-nicking enzymes. *Genome Res.* **2012**, *22*, 1327–1333.
83. Guilinger, J.P.; Thompson, D.B.; Liu, D.R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* **2014**, *32*, 577–582.
84. Ran, F.A.; Cong, L.; Yan, W.X.; Scott, D.A.; Gootenberg, J.S.; Kriz, A.J.; Zetsche, B.; Shalem, O.; Wu, X.; Makarova, K.S.; *et al.* *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* **2015**, *520*, 186–191.
85. González, F.; Zhu, Z.; Shi, Z.-D.; Lelli, K.; Verma, N.; Li, Q.V.; Huangfu, D. An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Stem Cell* **2014**, *15*, 215–226.
86. Moore, R.; Spinhirne, A.; Lai, M.J.; Preisser, S.; Li, Y.; Kang, T.; Bleris, L. CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells. *Nucleic Acids Res.* **2014**, *43*, 1297–1303.
87. Kim, S.; Kim, D.; Cho, S.W.; Kim, J.; Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* **2014**, *24*, 1012–1019.
88. Hendel, A.; Bak, R.O.; Clark, J.T.; Kennedy, A.B.; Ryan, D.E.; Roy, S.; Steinfeld, I.; Lunstad, B.D.; Kaiser, R.J.; Wilkens, A.B.; *et al.* Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **2015**, *33*, 985–989.
89. Fu, Y.; Sander, J.D.; Reyon, D.; Cascio, V.M.; Joung, J.K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **2014**, *32*, 279–284.
90. Cho, S.W.; Kim, S.; Kim, Y.; Kweon, J.; Kim, H.S.; Bae, S.; Kim, J.S. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **2014**, *24*, 132–141.
91. Koo, T.; Lee, J.; Kim, J.-S. Measuring and reducing off-target activities of programmable nucleases including CRISPR-Cas9. *Mol. Cells* **2015**, *38*, 475–481.
92. Gabriel, R.; Lombardo, A.; Arens, A.; Miller, J.C.; Genovese, P.; Kaepfel, C.; Nowrouzi, A.; Bartholomae, C.C.; Wang, J.; Friedman, G.; *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* **2011**, *29*, 816–823.
93. Wu, X.; Scott, D.A.; Kriz, A.J.; Chiu, A.C.; Hsu, P.D.; Dadon, D.B.; Cheng, A.W.; Trevino, A.E.; Konermann, S.; Chen, S.; *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* **2014**, *32*, 670–676.
94. Frock, R.L.; Hu, J.; Meyers, R.M.; Ho, Y.-J.; Kii, E.; Alt, F.W. Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* **2014**, *33*, 179–186.
95. Wang, X.; Wang, Y.; Wu, X.; Wang, J.; Wang, Y.; Qiu, Z.; Chang, T.; Huang, H.; Lin, R.-J.; Yee, J.K. Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat. Biotechnol.* **2015**, *33*, 175–178.
96. Crosetto, N.; Mitra, A.; Silva, M.J.; Bienko, M.; Dojer, N.; Wang, Q.; Karaca, E.; Chiarle, R.; Skrzypczak, M.; Ginalski, K.; *et al.* Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* **2013**, *10*, 361–365.
97. Tsai, S.Q.; Zheng, Z.; Nguyen, N.T.; Liebers, M.; Topkar, V.V.; Thapar, V.; Wyvekens, N.; Khayter, C.; Iafrate, A.J.; Le, L.P.; *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **2014**, *33*, 187–197.

98. Kim, D.; Bae, S.; Park, J.; Kim, E.; Kim, S.; Yu, H.R.; Hwang, J.; Kim, J.-I.; Kim, J.-S. Digenome-seq: Genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* **2015**, *12*, 237–243.
99. Canver, M.C.; Bauer, D.E.; Dass, A.; Yien, Y.Y.; Chung, J.; Masuda, T.; Maeda, T.; Paw, B.H.; Orkin, S.H. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J. Biol. Chem.* **2014**, *289*, 21312–21324.
100. Smith, C.; Gore, A.; Yan, W.; Abalde-Atristain, L.; Li, Z.; He, C.; Wang, Y.; Brodsky, R.A.; Zhang, K.; Cheng, L.; *et al.* Whole-Genome Sequencing Analysis Reveals High Specificity of CRISPR/Cas9 and TALEN-Based Genome Editing in Human iPSCs. *Stem Cell* **2014**, *15*, 12–13.
101. Suzuki, K.; Yu, C.; Qu, J.; Li, M.; Yao, X.; Yuan, T.; Goebel, A.; Tang, S.; Ren, R.; Aizawa, E.; *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*, 31–36.
102. Veres, A.; Gosis, B.S.; Ding, Q.; Collins, R.; Ragavendran, A.; Brand, H.; Erdin, S.; Cowan, C.A.; Talkowski, M.E.; 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*, 27–30.
103. Iyer, V.; Shen, B.; Zhang, W.; Hodgkins, A.; Keane, T.; Huang, X.; Skarnes, W.C. Off-target mutations are rare in Cas9-modified mice. *Nat. Methods* **2015**, *12*, doi:10.1038/nmeth.3408.
104. Jiang, J.; Lv, W.; Ye, X.; Wang, L.; Zhang, M.; Yang, H.; Okuka, M.; Zhou, C.; Zhang, X.; Liu, L.; *et al.* Zscan4 promotes genomic stability during reprogramming and dramatically improves the quality of iPS cells as demonstrated by tetraploid complementation. *Cell Res.* **2013**, *23*, 92–106.
105. Cradick, T.J.; Fine, E.J.; Antico, C.J.; Bao, G. CRISPR/Cas9 systems targeting-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **2013**, *41*, 9584–9592.
106. Yusa, K.; Rashid, S.T.; Strick-Marchand, H.; Varela, I.; Liu, P.-Q.; Paschon, D.E.; Miranda, E.; Ordóñez, A.; Hannan, N.R.F.; Rouhani, F.J.; *et al.* Targeted gene correction of  $\alpha$ 1-antitrypsin deficiency in induced pluripotent stem cells. *Nature* **2011**, *478*, 391–394.
107. Ousterout, D.G.; Perez-Pinera, P.; Thakore, P.I.; Kadi, A.M.; Brown, M.T.; Qin, X.; Fedrigo, O.; Mouly, V.; Tremblay, J.P.; Gersbach, C.A. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol. Ther.* **2013**, *21*, 1718–1726.
108. Tan, E.-P.; Li, Y.; Velasco-Herrera, M.D.C.; Yusa, K.; Bradley, A. Off-target assessment of CRISPR-Cas9 guiding RNAs in human iPS and mouse ES cells. *Genesis* **2015**, *53*, 225–236.
109. Karakikes, I.; Stillitano, F.; Nonnenmacher, M.; Tzimas, C.; Sanoudou, D.; Termglinchan, V.; Kong, C.-W.; Rushing, S.; Hansen, J.; Ceholski, D.; *et al.* Correction of human phospholamban R14del mutation associated with cardiomyopathy using targeted nucleases and combination therapy. *Nat. Commun.* **2015**, *6*, doi:10.1038/ncomms7955.