The FEBS Journal

FEBS PRESS science publishing by scientists

REVIEW ARTICLE

Potential pitfalls of CRISPR/Cas9-mediated genome editing

Rongxue Peng[1,2], Guigao Lin[1] and Jinming Li[1]

1 National Center for Clinical Laboratories, Beijing Hospital, Beijing, China
2 Graduate School, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China

Keywords
Cas9; CRISPR/Cas systems; DNA cleavage; gene targeting; genome editing; potential pitfalls; sgRNA; target specificity

Correspondence
J. Li, National Center for Clinical Laboratories, Beijing Hospital, No. 1 Dahua Road, Dongdan, Beijing 100730, China
Fax: +86 10 65212064
Tel: +86 10 58115061
E-mail: jmli@nccl.org.cn

(Received 15 June 2015, revised 25 September 2015, accepted 30 October 2015)

doi:10.1111/febs.13586

Recently, a novel technique named the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (Cas)9 system has been rapidly developed. This genome editing tool has improved our ability tremendously with respect to exploring the pathogenesis of diseases and correcting disease mutations, as well as phenotypes. With a short guide RNA, Cas9 can be precisely directed to target sites, and functions as an endonuclease to efficiently produce breaks in DNA double strands. Over the past 30 years, CRISPR has evolved from the 'curious sequences of unknown biological function' into a promising genome editing tool. As a result of the incessant development in the CRISPR/Cas9 system, Cas9 co-expressed with custom guide RNAs has been successfully used in a variety of cells and organisms. This genome editing technology can also be applied to synthetic biology, functional genomic screening, transcriptional modulation and gene therapy. However, although CRISPR/Cas9 has a broad range of action in science, there are several aspects that affect its efficiency and specificity, including Cas9 activity, target site selection and short guide RNA design, delivery methods, off-target effects and the incidence of homology-directed repair. In the present review, we highlight the factors that affect the utilization of CRISPR/Cas9, as well as possible strategies for handling any problems. Addressing these issues will allow us to take better advantage of this technique. In addition, we also review the history and rapid development of the CRISPR/Cas system from the time of its initial discovery in 2012.

Introduction

Subsequent to the discovery of the DNA double helix, studies have focused on manipulating the genomes of cells and organisms with the aim of exploring gene function. However, several of the earliest gene modification approaches, such as physical mutagenesis, chemical mutagenesis and transposon-mediated insertional mutagenesis, were not sequence-specific. In the 1970s, the endogenous mechanisms of DNA double-strand break repair was discovered [1]. Thus, methods for generating precise breaks at specific DNA sites

Abbreviations
AdVs, adenoviral vectors; BLESS, direct *in situ* breaks labeling, enrichment of streptavidin and next-generation sequencing; Cas protein, CRISPR-associated protein; ChIP-seq, ChIP and high-throughput sequencing; CRISPR, clustered regularly interspaced short palindromic repeat; crRNA, CRISPR RNA; dCas9, catalytically inactive Cas9; DNA-PKcs, DNA-dependent protein kinase catalytic subunit; GUIDE-seq, genome-wide, unbiased identification of DSBs enabled by sequencing; HDR, homology-directed repair; HR, homologous recombination; IDLV, integrase-defective lentiviral vector; NHEJ, nonhomologous end joining; NLS, nuclear location signal; PAM, protospacer adjacent motif; rAAV, recombinant adeno-associated viral vector; sgRNA, single guide RNA; TALEN, transcription activator-like effector nuclease; tracrRNA, trans-activating crRNA; ZFN, zinc finger nuclease.

were introduced as a valuable strategy for targeted genomic engineering.

One of the most significant breakthroughs was the development of gene targeting based on homologous recombination (HR) [2]. With HR-mediated targeting, it is possible to precisely manipulate any gene and establish knock-in and knockout animal models. However, despite such profound effects, the incidence of desired recombination events by HR-mediated targeting is extremely low and represents a time-consuming technology. To address this, a series of programmable nucleases, such as zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), have been developed in recent years for efficient and precise gene editing [3,4]. Despite having efficient gene editing properties, ZFNs and TALENs failed to be widely adopted as a result of their low-specificity and interference between contiguous modules in larger arrays.

To overcome these challenges, a novel RNA-guided endonuclease based genome editing technology named the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (Cas)9 system has been developed. With a short guide RNA (sgRNA), Cas9 can be guided to target sites precisely, and functions as an endonuclease to produce DNA double-strand breaks (DSBs) efficiently. We review the fundamental structures and the rapid development of the CRISPR-Cas9 system subsequent to it first being reported in 2012. We also discuss factors affecting its utilization, as well as possible strategies for handling any problems.

History of CRISPR/Cas9

The story of CRISPR began in 1987. When Ishino *et al.* [5] were studying the mechanism underlying the isozyme conversion of alkaline phosphatase in *Escherichia coli*, they discovered several 'curious sequences' in the 3′-end flanking region of the *iap* gene and described it as a set of 29 nucleotide repeats with 32 nucleotide spacing sequences. Over the next decade, many other repeat elements were found in numerous bacteria and archaea. According to Mojica *et al.* [6], these short regularly spaced repeats were present in more than 40% of bacteria and 90% of archaea.

In 2002, these short repeats were officially named Clustered Regularly Interspaced Short Palindromic Repeats, and the abbreviation CRISPR began to circulate widely [7]. In the next few years, a series of *Cas* genes, situated next to the CRISPR locus, were identified in CRISPR-containing prokaryotes [8]. Although several CRISPR loci have been discovered in numerous microbial species, their biological significance remained a puzzle until 2005. At that time, three independent research groups reported that CRISPR contains spacers of extra chromosomal origin, including spacers from phages and plasmids [6,9]. In 2007, after further analysis and comparison, it was confirmed that CRISPR systems were a kind of bacterial adaptive immune defense mechanism that protected bacteria against plasmids and phages [10].

Subsequently, with the discovery of the *Cas* gene, Cas protein, protospacer adjacent motif (PAM), CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA), many details regarding the CRISPR systems were quickly unveiled [9,11,12]. Although the mechanisms behind CRISPR systems have not been determined completely, their basic functions and processes were becoming clear. In addition, the classification of CRISPR systems was also clarified. Based on the diverse homology of Cas proteins, CRISPR systems were classified into three categories: type I, type II and type III. The type II system was found to need only one Cas protein to recognize and cleave target sites, whereas type I and type III CRISPR systems required a set of Cas proteins [11,13,14]. The unique protein of the type II CRISPR system is Cas9, which was confirmed to be an essential component of the CRISPR systems [13].

In view of the function and mechanism of type II CRISPR systems, studies began to explore its application in genome editing. A key insight came in 2012 with the observation that Cas9-crRNA complexes of the *Streptococcus thermophilus* and *Streptococcus pyogenes* could function as RNA-guided endonucleases *in vitro*, with RNA recognizing target sequences and Cas9 protein cleaving DNA at specific sites [15]. These findings, along with previous studies, led to the proposal that the Cas9-crRNA complex could be a robust genome editing tool for inducing specific double-strand breaks. Subsequently, the CRISPR/Cas system has come of age as a novel targeted genome engineering technology and has been successfully used in numerous species. Furthermore, the technique is constantly being modified and optimized with the aim of achieveing different outcomes, and new studies are continually forthcoming. During the past 30 years, CRISPR has evolved from 'curious sequences of unknown biological function' into a promising genome editing tool that is employed globally.

Type II CRISPR/Cas9 for genome editing

The type II system genomic CRISPR locus includes the *tracrRNA* gene, *Cas* gene and CRISPR

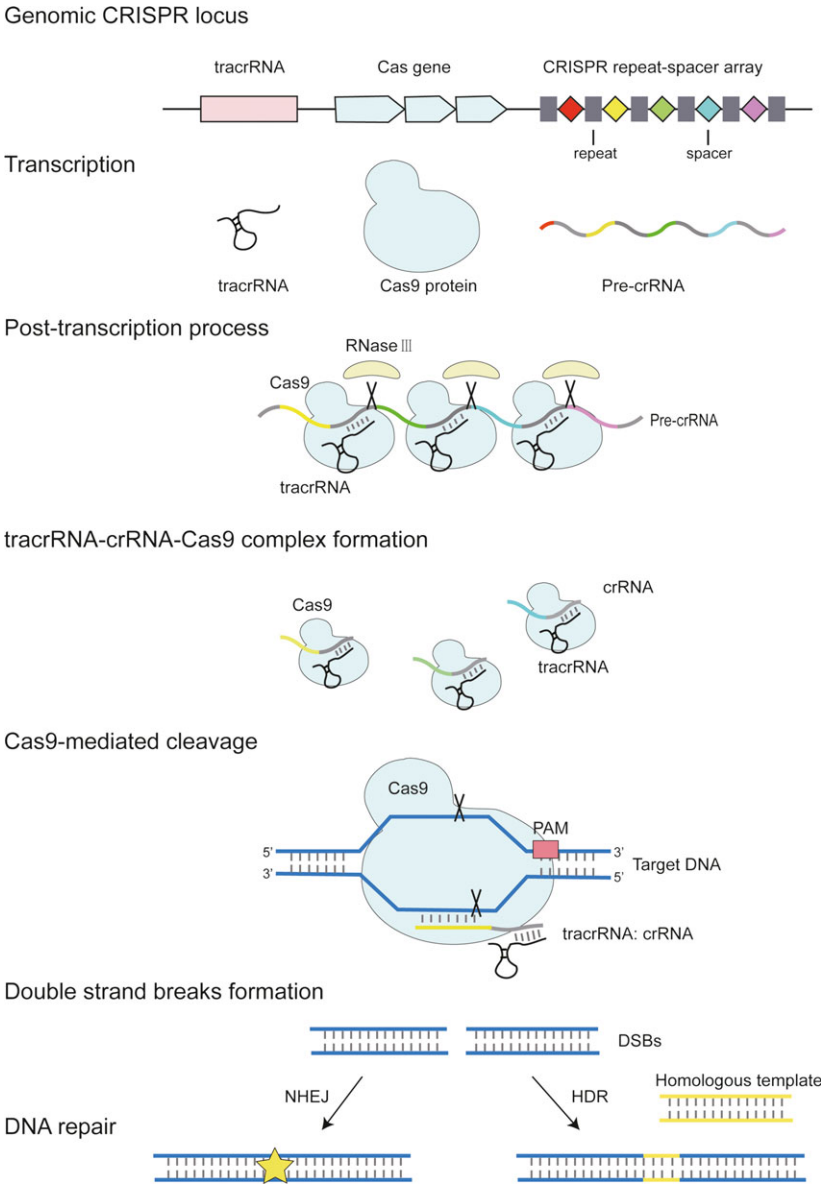

Fig. 1. Biology of the type II CRISPR/Cas system. Using the *Streptococcus pyogenes* type II CRISPR/Cas system as an example, the genomic CRISPR locus is transcribed into tracrRNA, Cas9 and pre-crRNA. With the cooperation of tracrRNA and RNase III, pre-crRNA trims further, forming the tracrRNA:crRNA duplex, which directs Cas9 to generate DSBs at target sites. Finally, DSBs can be repaired through the NHEJ or the HDR pathway.

repeat-spacer array, which are transcribed into tracrRNA, Cas9 proteins and pre-crRNA, respectively (Fig. 1). With the cooperation of tracrRNA and RNaseIII, pre-crRNA can be cut into crRNAs that can interact with tracrRNAs and subsequently lead Cas9 to recognize the specific DNA sites [16]. The Cas9:RNA complex randomly searches DNA sequences and rapidly dissociates from non-PAM sites. Only by binding to a PAM sequence, a short motif adjacent to the target sequence (usually NGG motif for SpCas9, sometimes NAG motif), does the Cas9:RNA complex interrogate the flanking DNA sequences for gRNA complementarity [17]. At the target sites that match the tracrRNA:crRNA duplex and are flanked by PAMs, the HNH nuclease domain of Cas9 cleaves the strand that binds to crRNA, and the RuvC-like domain cuts the other DNA strand, forming DSBs at the specific sites [12].

Once the site-specific DSBs are created, two different repair mechanisms can be activated: the non-homologous end joining (NHEJ) and the homology-directed repair (HDR) mechanisms. When there are no templates, NHEJ can connect the break sequences directly and induce insertions or deletions (indels). If indels cause a frame shift in the reading frame, this can result in the creation of premature stop codons [18]; if insertions or deletions comprise three bases (or multiples of three bases) they may affect protein

function [19]. In the presence of a donor template, the HDR pathway is activated to induce specific insertions, deletions or mutations [20].

As a result of its simplicity, the type II CRISPR/Cas system has been developed as a robust programmable tool, which is known as CRISPR/Cas9. After Jinek *et al.* [15] first utilized CRISPR/Cas9 to induce specific DSBs, Mali *et al.* [21] began to synthesize an artificial Cas9 protein and cloned it into a mammalian expression system. In addition, they expressed a chimeric sgRNA to replace the tracrRNA:crRNA duplex (Fig. 2). Using a custom CRISPR system, desired mutations can be introduced in several human cell lines, with targeting rates ranging from 2% to 38%. Meanwhile, Cong *et al.* [22] designed two different type II CRISPR/Cas systems and demonstrated that Cas9 nucleases can induce precise cleavages at specific sites under the guidance of short RNAs in human and mouse cells.

With the gradual perfection of CRISPR/Cas9 systems, *S. pyogenes* Cas9 (SpCas9) co-expressed with custom guide RNAs (sgRNAs or tracrRNA:crRNA duplexes) have been successfully used in bacteria, fungi, viruses, parasites, plants, animals and human cell lines [18,23,24]. Targeting with multiple sgRNAs was also successfully applied to multiplex genome engineering [25]. The 'humanized' CRISPR/Cas9 system has overturned previous methods of animal model generation, functional genomic screens, transcriptional modulation, epigenetic control and live imaging of the cellular genome. However, although CRISPR/Cas9 has a broad range of action, several aspects still require further investigation and there are some points that remain worthy of attention.

Key factors affecting the CRISPR/Cas9 system

To date, the CRISPR/Cas9 system has already shown itself to comprise a robust and flexible tool for genome editing and gene regulation. With further research on CRISPR, however, it became apparent that this technology was not as easy as once assumed. A large number of studies have investigated diverse factors affecting the CRISPR/Cas9 system, such as Cas9 activity, target site selection and sgRNA design, delivery methods, off-target effects, and the incidence of HDR (Fig. 3). By addressing these potential pitfalls, we can take better advantage of this technique, as well as improve its efficiency and specificity.

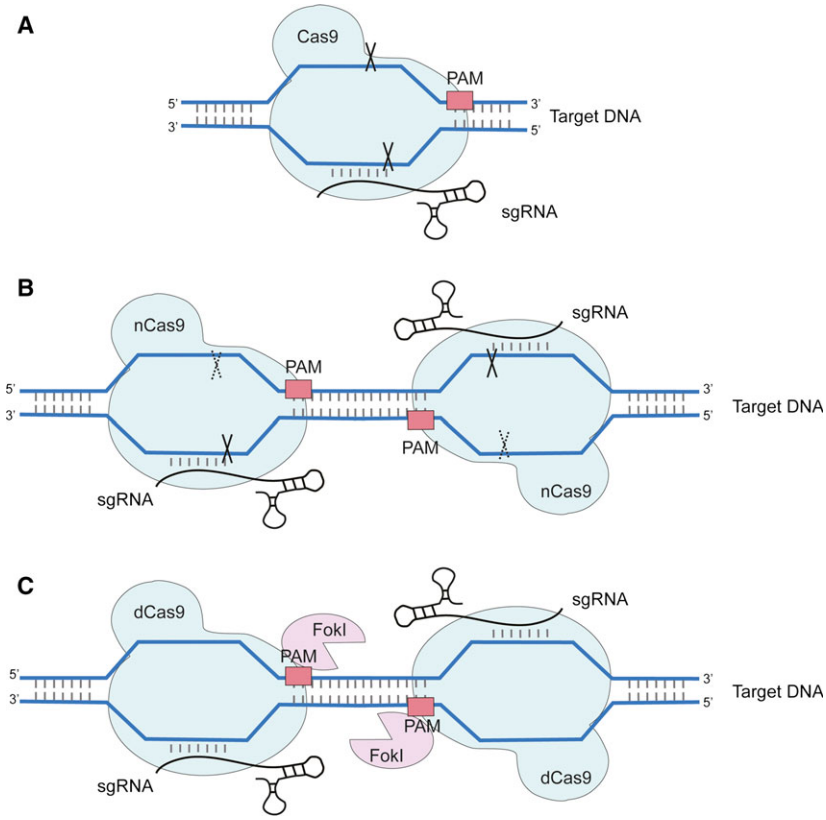

Fig. 2. CRISPR/Cas9 for genome editing. (A) With the guidance of sgRNA, the HNH nuclease domain of Cas9 cleaves the strand that binds to sgRNA and the RuvC-like domain cuts another DNA strand, forming DSBs at target sites. (B) Paired Cas9 nickases (i.e. mutating either the HNH or the RuvC-like domain in Cas9 generates a variant protein with single-strand DNA cleavage capacity) and sgRNAs are used to avoid off-target effects. (C) Dimeric sgRNA-guided fCas9 (dCas9 fused to FokI nuclease domains) are used to improve the specificity of genome editing.

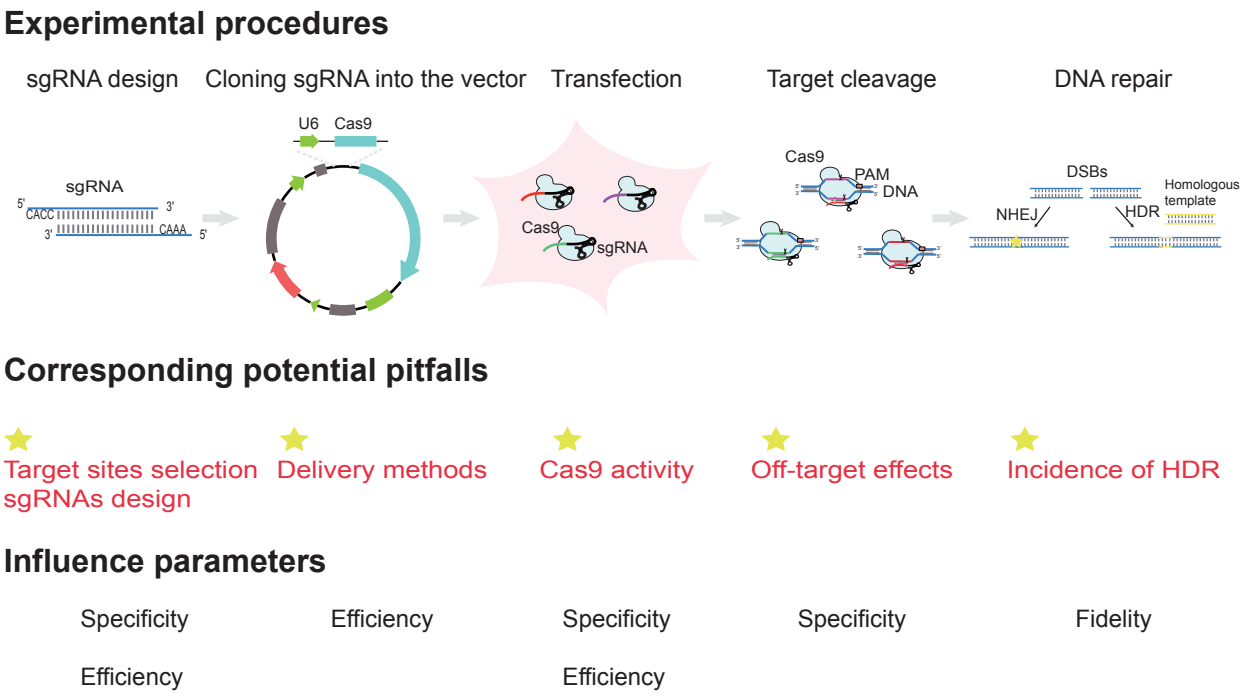

Fig. 3. Overview of experiments and potential pitfalls of the CRISPR/Cas9 system. Custom sgRNAs for each target are designed by hand or using appropriate software. sgRNA can be cloned into the vector for co-expression with Cas9. Completed vectors and homologous templates are then delivered into target cells. Finally, the target DNA can be cleaved specifically. During this entire process, diverse potential pitfalls have been found, such as Cas9 activity, target site selection and sgRNA design, delivery methods, off target effects and the incidence of HDR, which might affect the efficiency and specificity of the CRISPR/Cas9 system. Details of potential pitfalls and possible strategies for addressing these factors are provided in the main text.

Cas9 activity

As a genome editing tool, the CRISPR/Cas9 system cleaves specific nucleotides based on sequence complementarity with only two significant components: the Cas9 protein and sgRNA. With the binding of sgRNA and target DNA, the Cas9 protein undergoes a large-scale conformational rearrangement. Accordingly, the catalytic nuclease lobe of Cas9 rotates ~ 100°, generating nucleic acid-cleaving activity [26].

Generally, Cas9 can recognize genomic loci under the guidance of sgRNAs that bind to 20 nucleotide target sequences. However, Hsu *et al.* [27] found that sgRNAs with +85 nucleotide tracrRNA tails increased the activity of Cas9 and induced higher level of indels *in vivo*. They also observed that both concatenated and interspaced two base mismatches, which occurred in the proximal region of PAM, greatly reduced Cas9 activity. This effect was further expanded to three concatenated mismatches [27], and three or more interspaced and five concatenated mismatches were found to abrogate Cas9 cleavage activity in most genes [27]. Additionally, the results of further investigations suggested that excessively truncated guide RNA would also result in Cas9 losing cutting activity [28]. Therefore, to provide higher Cas9 cutting efficiency, optimization of the sgRNA design and cautious selection of target sites are badly needed. In addition, increasing exogenous sgRNA was also found to improve DNA cleavage activity [29]. In theory, more sgRNA:Cas9 complexes can promote higher editing efficiency. However, excessive sgRNA:Cas9 complexes may give rise to off-target effects as a result of the inevitable complementarity of nonspecific sequences in the genome [30]. An up to six-fold molar excess of sgRNA over Cas9 protein has been shown to maximize the on-target mutation frequencies when the recombinant Cas9 protein was delivered into cultured human cells with *in vitro* transcribed sgRNA [31]. Thus, to improve on-target mutation rates, as well as the activity of Cas9, the concentrations of sgRNA and Cas9 should be considered.

The activity of Cas9 protein can also be affected by other elements to a greater exstent than by sgRNAs. In eukaryotic gene editing, Cas9 is always connected with the nuclear location signal (NLS) to translocate into the nucleus. Shen *et al.* [32] demonstrated that

adding a linker containing 32 amino acids between Cas9 and NLS could enhance its DNA cleavage activity. In addition, several evolutionarily divergent Cas9 proteins have been used for genome engineering, such as *Neisseria meningitidis* Cas9 (NmCas9) [33], *S. thermophilus* Cas9 (St1Cas9) [34] and *Staphylococcus aureus* Cas9 (SaCas9) [35]. Compared to typical SpCas9, most of these orthologous Cas proteins have different PAM recognition sequences and variable activity. In human cells transfected with plasmids, St1Cas9 displayed high activity and SaCas9 showed greater efficiency, as did SpCas9 [34]. In lentivirus-mediated transduction, however, St1Cas9 displayed a lower than average mutation rate compared to that of SpCas9 [36], indicating that the selection of a specific orthologonal Cas9 should be considered on a case-by-case basis.

Target site selection and sgRNA design

Among the potential pitfalls of CRISPR/Cas9 systems, sgRNA design is a prime concern. Because CRISPR/Cas9 systems are highly programmable, Cas9/sgRNA complexes can be utilized for genome editing or catalytically inactive Cas9 (dCas9)/sgRNA complexes can be used for gene regulation. These applications require the design of sgRNAs that are efficient and specific. However, because this requires the consideration of many criteria, rational sgRNA design remains a major challenge.

Previously, it was assumed that Cas9/sgRNA complexes could cleave double-strand DNA in the presence of PAM and an adjacent complementary target sequence. However, many experiments showed that some sgRNAs were less efficient or even inactive [37–40, 41]. For genome editing experiments, a pool of sgRNAs first required to be screened for activity; hence, design criteria to maximize sgRNA efficiency are a valuable pursuit. With accumulating experimental data on the use of CRISPR/Cas9 systems for genome engineering, a range of sequence features in and around the target sequences that predict sgRNA efficiency have been identified.

For CRISPR/Cas9-mediated genome editing, the 5′ end of sgRNAs that append a G (guanine) (e.g. GX19NGG) is strongly needed for expression from a U6 promoter [37]. In addition, G is preferred in the first or second position closest to PAM, which may assist Cas9 loading [37], whereas C (cytosine) is strongly unfavorable in the same positions. Third, because multiple U (uracil) in sgRNA cause low sgRNA expression, T (thymine) is undesirable at the four nucleotide positions adjacent to PAM [38]. A (adenine) is preferred in the middle of sgRNA and G is preferred in the PAM-distal region [37]. In all, G-rich and A-depleted sgRNAs are more stable and more efficient [39]. Moreover, novel features of the SpCas9 PAM that reproducibly impact upon sgRNA activity have been found. For example, there is a preference in the variable nucleotide of NGG, where C is favored and T is disfavored [40]. An extended PAM sequence of CGGH is optimal for the use of SpCas9 to generate DSBs in mammalian cells. Conversely, TGGG shows the lowest activity [40]. Compared to the sequence features in the CRISPR/Cas9-mediated genome editing, the sequence preference for dCas9 fusion-mediated inhibition/activation (CRISPRi/a) is substantially different. In CRISPRi/a experiments, 19 nucleotide sgRNAs have the highest efficiency, and perform better than the truncated sgRNAs with 17–18 nucleotide spacers and the sgRNAs with 20 nucleotide spacers [38]. Additionally, purines are preferred at most positions in sgRNAs [38]. By contrast to the 'seed sequences' that dominate sequence preferences in the CRISPR/Cas9 system, most of the nucleotides in the spacer region jointly contribute to sgRNA efficiency in the CRISPR/dCas9 system [38].

With so many criteria proposed, a growing number of computational tools now facilitate the design of sgRNAs. Most of the current sgRNA design tools support either the SpCas9 system or multiple orthogonal Cas9 systems from other bacterial species. sgRNA design software that enables a comparison between them is listed in Table 1.

Delivery methods

CRISPR/Cas9 technology is changing the field of genome engineering and is also expected to change the treatment of genetic diseases. Achieving this goal requires not only improvements in efficacy and specificity, but also the optimization of delivery methods.

The introduction of plasmids that simultaneously encode sgRNA and Cas9 into target cells by electroporation, nucleofection or lipofectamine represents a common and rapid method that can be applied to a wide variety of cell lines. The most commonly used plasmids always express an optimized SpCas9 and chimeric gRNA. With the demand for large fragment expression, multiple plasmids have been used to target different sites [25]. However, all or part of the plasmids are often randomly integrated into the host genome [31]. Plasmid DNA can also be inserted into both on-target and off-target sites, which can lead to difficulties in detection. Furthermore, host immune

Table 1. Comparison of different web-based software for sgRNA design.

Tool	Website	Type of CRISPR/Cas system	Species support	CRISPR/Cas designs		Off-target analysis	Batch mode	Comments	Citation
				Nucleases	Nickase				
ZiFiT	http://zifit.partners.org/ZiFiT	Type II only	9	Yes	Yes	Yes	No	Analysis DNA sequence up to 1 kb; one of the earliest tools for searching potential target sites; this software also be applied in ZFNs design and TALENs design	[74]
CRISPR design	http://crispr.mit.edu	Type II only	16	Yes	Yes	Yes	Yes	Analysis DNA sequence up to 500 bp; display the off-target scores in an unambiguous manner; limited exclusively to 'NGG' PAM sequences	[27]
CRISPR direct	http://crispr.dbcls.jp	Type II only	18	Yes	No	Yes	No	Identify various PAM sequences; display main sequence features of the candidate sites	[75]
CRISPR RGEN tools	http://www.rgenome.net	Different Type II	16	Yes	No	Yes	No	Analysis DNA sequence up to 1 kb; standalone version available; random mismatches can be searched	[76]
CHOPCHOP	https://chopchop.rc.fas.harvard.edu	Different Type II	25	Yes	No	Yes	No	Display in dynamic graphical interface; design sgRNAs for orthogonal Cas9 with different PAM sequences; this software also be applied in TALEN design	[77]
E-CRISPR	http://www.e-crisp.org/E-CRISP	Different Type II	33	Yes	Yes	Yes	No	Start application and design purpose can be set for specific experimental goals (e.g. knockout, N-terminal tagging, C-terminal tagging, CRISPRi/a); enable gene annotation filtering	[78]
sgRNA Designer	http://broadinstitute.org/rnai/public/analysis-tools/sgrna-design	Type II only	2	Yes	No	No	Yes	Analysis DNA sequences up to 10 kb; standalone version available; for human and mouse genes or interest gene only; use experimentally defined scoring scheme	[31,40]
CRISPR MultiTargeter	http://www.multicrispr.net	Multiple types	12	Yes	Yes	No	Yes	Enable multiple CRISPR systems sgRNA design; scan off-target sites with the aid of GT-scan or Cas OFFinder; basic sgRNA target sites search in a set of similar genes or transcripts	[79]

Table 1. (Continued).

Tool	Website	Type of CRISPR/Cas system	Species support	CRISPR/Cas designs		Off-target analysis	Batch mode	Comments	Citation
				Nucleases	Nickase				
CRISPR-ERA	http://crispr-era.stanford.edu/InitAction.action	Type II only	9	Yes	Yes	Yes	No	Design sgRNAs for gene regulation; use E-score to reveal the sequence features and S-score to display off-target effects	[80]
sgRNA Scorer	https://crispr.med.harvard.edu/sgRNAScorer	Different Type II	12	Yes	No	No	No	Analysis DNA sequences up to 10 kb; scan off-target sites with the aid of Cas OFFinder; use a new *in vivo* and multiplex library-on-library methodology to assess sgRNA activity across ~ 1400 genes	[36]
CRISPRscan	http://crisprscan.org	Type II only	7	Yes	No	Yes	No	Provide a valuable resource for predicting the most efficient sgRNA for *in vivo* targeting, especially zebrafish; commonly criteria to design sgRNAs are outclassed	[39]

responses that depend on the recognition of unmethylated CpG dinucleotides in the bacterial DNA can be induced by these inserted bacterial sequences and may interrupt the process of genome editing [42]. In addition, these methods of transfection are inefficient in primary cells and may lead to cytotoxicity. Once transfected, plasmid DNA can also persist inside the cells for several days, which may aggravate off-target effects [43]. Therefore, other approaches, including the use of recombinant proteins, *in vitro* transcribed mRNA and viral vectors, are being investigated.

To circumvent these challenges in the use of plasmids, the recombinant Cas9 protein complex with sgRNAs has been delivered into cultured human cells, mice and zebrafish via electroporation or microinjection [31]. In addition, *in vitro* transcribed Cas9 mRNA has been microinjected into zebrafish, mice, pigs and monkeys to generate genetically modified animals. However, these methods may be stressful to cells. Alternative approaches including the chemical conjugation of Cas9 protein and sgRNAs, such as cell-penetrating peptide-conjugated recombinant Cas9 protein [44], cationic lipid mediated delivery of Cas9 protein [45] and delivery of self-assembled DNA nanoclew-mediated Cas9/sgRNA complexes [46], were applicated gradually. By contrast to Cas9 mRNA, which first needs to be translated after transfection into cells, Cas9 protein performs immediately upon transfection [31]. In addition, because translation begins only at the two-cell stage, injection of mRNA encoding cas9 into single-cell embryos often results in chimerism [47]. Using Cas9 protein can circumvent this limitation. Moreover, the detection of Cas9 protein is much easier than the detection of Cas9 mRNA before transfection. Although these methods of delivery can reduce off-target effects relative to plasmid-mediated delivery, the efficiency is still considered insufficient.

By contrast to microinjection, chemical transfection and electroporation, gene transfer by viral vectors is an active, receptor-dependent process that allows control over the amount of DNA transfected for greater efficiency and less cytotoxicity. Moreover, viral vectors can enter into large numbers of cell types efficiently, both *in vivo* and *in vitro*. Therefore, viral vectors are widely used to deliver CRISPR systems. To date, integrase-defective lentiviral vectors (IDLVs) [48], adenoviral vectors (AdVs) [49] and recombinant adeno-associated viral vectors (rAAVs) [18] have been used to transfer CRISPR systems into mammalian cells. Given their large packaging capacity (~ 10 kb), IDLVs comprise a useful tool for the delivery of CRISPR/Cas9 systems. They have been used for the construction of large-scale sgRNA expressing arrays containing

a vast number of human and mouse genes [50], and have facilitated the detection of off-target effects with low frequency (1%) [48]. Although the transient expression of IDLVs-delivered Cas9 has an advantage in terms of safety in rapidly dividing cells, Cas9 expression may persist after IDLV delivery in quiescent and slowly dividing cells, leading to high off-target effects. Similar to IDLVs, AdVs are also commonly used viral vectors for the delivery of CRISPR/Cas9 components. However, the difference between these vectors is that the linear double-stranded AdV DNA has a terminal protein at its 5′ ends as a result of covalent binding [49]. This protein-capped DNA structure reduces the possibility for interactions between the exogenous DNA and off-target sites, making chromosomal integration less promiscuous and targeted insertion more precise [49]. Finally, because of their nonpathogenicity, low immunogenicity, nonintegrating nature and wide range of serotype specificity, rAAV vectors have also been deployed to introduce CRISPR components into animals with the aim of investigating gene function and establishing desired animal models [18]. Because rAAV only leaves ~ 4.5 kb for packaging CRISPR/Cas9 components, using a single rAAV vector to deliver the typical SpCas9 (~ 4.2 kb), chimeric sgRNA and control elements still represents a substantial challenge [51]. Alternative approaches are the use of two separate rAAV vectors, and the delivery of a rationally-designed truncated SpCas9. However, both methods have demonstrated reduced activity [35]. More recently, a smaller Cas9 from *S. aureus* has been identified that resolves such problems. Ran *et al.* [35] achieved > 40% gene modification efficiency of the *PCSK9* gene in murine liver by intravenous injection of a rAAV vector that packaged SaCas9 and sgRNA [35]. Nevertheless, rAAV-mediated delivery requires further optimization before it can be used for therapeutic applications.

Off-target effects

Many studies have shown that the CRISPR/Cas9 system is a simple but highly efficient approach for genome editing in a variety of cells and organisms, both *in vitro* and *in vivo*. The Cas9 nuclease can be directed via complementary base pairing between the first 20 nucleotides of sgRNA and target DNA sequences that lie adjacent to PAM. With the cleavage of Cas9, target sequences generate DSBs and introduce new genetic materials. Nevertheless, a number of studies have demonstrated that the CRISPR/Cas9 system can induce a substantial amount of off-target mutagenesis [30,52]. These off-target effects might play a role in recognizing and destroying hypervariable viral nucleic acids or plasmid DNA, which is beneficial to bacteria and archaea [53]. However, for biological studies and genetic therapies, off-target phenomena generate undesired mutations at random sites, thus impacting precise gene modification.

Previous studies have shown that Cas9-mediated cleavage can be inhibited by a single mismatch in the complementary region of sgRNA and target sequences, especially the PAM-proximal nucleotides [30]. On the basis of whole-genome sequencing, several studies have shown that mismatches at the 5′-terminal of the target sites are generally better tolerated [22,54]. Moreover, studies in human cells have shown that up to five mismatches will not prevent cutting of target sequences, and that off-target sites can be mutagenized at frequencies higher than the intended on-target sites [30]. Thus, for a better application of the CRISPR/Cas9 system, it is important to evaluate the potential hybrid effects of high-frequency off-target mutations.

A series of experimental and computational methods that predict off-target sites were discovered based on modification of Cas9 *in vivo* [27,52]. Initially, efforts to depict the specificity of Cas9 cleavage focused exclusively on the SpCas9 nuclease. Off-target sites were deduced from alignment algorithms using a 'seed sequence' adjacent to the NGG PAM. When NAG PAM sequences also recognized by SpCas9 were recognized, it was apparent that many off-target mutations had not been identified. Subsequently, a number of studies used sequence similarity-based off-target searches or dCas9-mediated ChIP to detect potential off-target sites for mutations [55]. However, such approaches are biased and not comprehensive. With the development of high-throughput sequencing, several new approaches that can reveal genome-wide off-target binding in an unbiased manner have been developed; for example, IDLV capture [48]; genome-wide unbiased identification of DSBs enabled by sequencing (GUIDE-seq) [56]; Digenome-seq [57] and direct *in situ* breaks labeling, enrichment of streptavidin, and next-generation sequencing (BLESS) [35]. Each of these methods has both advantages and drawbacks. IDLV capture uses viral genomes to tag the DNA DSBs, although this method can only detect off-target sites occurring at a frequency of > 1% [48]. Although GUIDE-seq and Digenome-seq can identify off-target sites that are cleaved at frequencies of < 0.1%, the former cannot work for DNA nicks and the latter does not take into account cellular factors that may affect off-target cleavages [56,57]. Unlike GUIDE-seq and IDLV capture, there is no need to

introduce exogenous bait for BLESS [35]. Moreover, it can identify not only the off-target sites previously predicted by sequence similarity-based off-target searches or ChIP, but also the off-target sites induced by insertion or deletion mismatches in the gRNA:DNA heteroduplex [35]. Although they are comprehensive and unbiased, these methods still need to be refined because the off-target effects identified *in vitro* might be different from those *in vivo*.

We are now on the threshold of understanding the specificity of CRISPR/Cas9 systems. Over the past 2 years, considerable efforts have been made to diminish off-target effects, and the specificity of CRISPR/Cas9 has been enhanced using the strategies outlined below.

(a) Selection of proper target sites and rational design of highly active sgRNAs. It has been reported that target site sequences with relatively low GC contents ($\leq 35\%$) are less likely to induce off-target effects, whereas high GC-content makes RNA/DNA hybrids more stable and increases the tolerance of mismatches [55]. Moreover, target sequences and PAM sequences containing at least three mismatches that form DNA bulges at the 5′-terminal, 3′-terminal or 7–10 bp adjacent to PAM should be avoided [55].
(b) Use paired Cas9 nickase and sgRNAs. Cas9 enzyme can be converted into Cas9 nickase, a variant protein with single-strand DNA cleavage capacity, by mutating either the HNH or the RuvC-like domain. With a pair of sgRNAs, Cas9 nickase can induce the generation of two closely adjacent single-strand nicks, which will eventually form DSBs [58] (Fig. 2).
(c) Use truncated sgRNAs. Fu *et al.* [59] demonstrated that using truncated sgRNAs with a length of 17–18 nucleotides can decrease undesired off-target effects by more than 5000-fold, as well as retain on-target efficiency [59]. Furthermore, combining truncated sgRNAs with paired Cas9 nickases can lead to further reductions in off-target mutations [59].
(d) Use dimeric sgRNA-guided fCas9 (dCas9 fused to FokI nuclease domains) (Fig. 2). Similar to paired Cas9 nickase, dCas9-FokI dimers can edit endogenous genes by constituting a functional FokI endonuclease [60]. fCas9 has been shown to modify on-target DNA sites with a specificity that is 140-fold higher than for typical Cas9 in human cells [61]. Moreover, fCas9 generally induces less off-target effects than paired Cas9 nickase [62].

Although these strategies greatly improve the specificity of on-target cleavage, they also demonstrate many defects, such as larger protein sizes, the need for multiple adjacent sgRNAs, incomplete avoidance of off-target effects or a decrease in on-target efficiency. Further improvements in the rational design of sgRNAs, optimization of delivery methods or the characterization of novel orthologous Cas proteins will aid future developments.

Incidence of HDR

As noted previously, DSBs cut by Cas9 nuclease at desired target sites can stimulate two distinct endogenous DNA repair mechanisms, NHEJ and HDR, to introduce loss of function mutations or precise modifications. NHEJ is initiated by the recruitment of Ku70/80 heterodimer and DNA-dependent protein kinase catalytic subunit (DNA-PKcs) [63]. This complex stabilizes the two DNA ends and recruits a series of proteins, including DNA ligase, to ligate the DNA breaks [64]. In HDR, the DNA DSBs are resected to generate 3′-ssDNA overhangs that can recruit a set of repair proteins and invade with a homologous donor template [64]. In general, even in the presence of donor templates, NHEJ is the more frequently repair pathway when using CRISPR/Cas9 systems [65]. Because NHEJ is error-prone but highly efficient, it is suitable for generating indels. However, it cannot edit the genome precisely or in the same way as HDR. As a result of the low incidence of HDR, improving the incidence of precise insertional mutagenesis using CRISPR/Cas9 is a challenge.

Cell lines deficient in NHEJ components increase levels of HDR. This suggests that these two repair processes are intensely competitive. Hence, suppressing key NHEJ molecules, such as KU70, KU80 or DNA ligase IV, by gene silencing [20], microinjection of ssRNAs [66] or the use of small molecule inhibitors [20,63,64,67–69] has been shown to increase the frequency of repair by the Cas9-mediated HDR pathway. One of the most prominent research achievements in the field has been the use of the inhibitor Scr7, which targets DNA ligase IV, to increase the efficiency of HDR-mediated genome editing up to 19-fold [20,68,69]. Moreover, from a collection of ~ 4000 small molecules of known biological activity, Yu *et al.* [67] identified two small molecules, L755507 and Brefeldin A, that can increase the HDR efficiency by two- to three-fold. Robert *et al.* [64] also confirmed that two small molecule inhibitors of DNA-PKcs (NU7441 and KU-0060648) could reduce the incidence of NHEJ and increase the frequency of HDR. In addition, short hairpin RNA sequences to knockdown KU70/80 or DNA ligase IV have also been used to promote the

efficiency of HDR in both human and mouse cells [20].

Although the suppression of key NHEJ molecules, KU70, KU80 or DNA ligase IV, can stimulate HDR, which is followed by Cas9 cleavage, these inhibitors may have toxic effects. Thus, a second strategy utilizing cell cycle synchronization techniques has been developed. Generally, the efficiency of Cas9-mediated gene editing in mice via NHEJ can reach 20–60%, whereas the efficiency of HDR is only 0.5–20% [65]. Although cells have different abilities with respect to repairing DSBs using either NHEJ or HDR, the phase of the cell cycle also regulates repair pathway selection. NHEJ always occurs during the entire cell cycle, whereas HDR is restricted to the late S and G2 phases [41]. Therefore, a new strategy has been developed that combines well-established cell cycle synchronization techniques with direct nucleofection of pre-assembled Cas9 ribonucleoprotein complexes to achieve controlled nuclease action at the optimal phase of the cell cycle for HDR [41].

There has been some progress in improving the efficiency of HDR, although very few studies are currently being published in this area. Whether for mass-producing animal models or for prospective gene therapy, precise genome editing is critical. Therefore, significant efforts should be made to enhance the incidence of HDR.

Conclusions and prospects

Systematic analyses of gene function have been hampered by the lack of suitable tools for precise and efficient gene engineering. The CRISPR/Cas9 system, which simply uses a guide RNA and Cas9 nuclease to identify and cut target DNA sequences, comprises a robust technology that has been used in diverse and innovative applications in biology. It has incomparable advantages over other gene editing tools. For example, the CRISPR/Cas9 system has more target sites than ZFNs and TALENs, and Cas9 has many variants that can be used in a variety of studies. Moreover, the system is extremely easy to use and can be executed in almost any laboratory. Cas9-based tools have greatly enhanced our ability to perform systematic analyses of gene function, as well as to reproduce tumor-associated chromosomal translocations precisely. This technology has also paved the way for the dissection of redundant gene functions, epigenetics and possible gene therapy.

Following the first publication of CRISPR/Cas9, countless studies have been conducted and many breakthroughs have been achieved in this field [70–72]. Despite the many advantages of this system, there are some challenges to the current Cas9-based tools, such as delivery methods, off-target effects and the balance of HDR/NHEJ pathways. In addition, some of the fundamental attributes of the system remain unclear, including the catalytic mechanism of Cas9, the mechanisms of underlying target sites identification and the basis for PAM-dependence. An understanding of these aspects will aid in efforts aiming to improve its catalytic efficiency, expand our choices of target sites and generate highly specific Cas9-mediated tools.

With the increasing application of CRISPR/Cas9, this system has been applied to gene editing in human tripronuclear zygotes [73], generating a tremendous amount of controversy. Undoubtedly, the results of such a study emphasize that the translation of CRISPR technologies from basic research to clinical practice is still a distant prospect. Further optimization and improvement of Cas9-based tools for genome editing is needed to meet the requirements of gene therapy application.

Acknowledgements

This work was supported by the capital development fund from the Beijing Municipal Public Health Bureau (No. 2011-4011-02) and the Special Fund for Health-Scientific Research in the Public Interest from the National Population and Family Planning Commission of the People's Republic of China (No. 201402018).

Author contributions

RP wrote the manuscript.

References

1 Scherer S & Davis RW (1979) Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proc Natl Acad Sci USA* **76**, 4951–4955.
2 Capecchi MR (1989) Altering the genome by homologous recombination. *Science* **244**, 1288–1292.
3 Bibikova M, Golic M, Golic KG & Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169–1175.
4 Moscou MJ & Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501.
5 Ishino Y, Shinagawa H, Makino K, Amemura M & Nakata A (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion

in *Escherichia coli*, and identification of the gene product. *J Bacteriol* **169**, 5429–5433.
6 Mojica FJ, Diez-Villasenor C, Garcia-Martinez J & Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**, 174–182.
7 Jansen R, Embden JD, Gaastra W & Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565–1575.
8 Schouls LM, Reulen S, Duim B, Wagenaar JA, Willems RJ, Dingle KE, Colles FM & Van Embden JD (2003) Comparative genotyping of Campylobacter jejuni by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. *J Clin Microbiol* **41**, 15–26.
9 Bolotin A, Quinquis B, Sorokin A & Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561.
10 Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA & Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712.
11 Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV & van der Oost J (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964.
12 Haurwitz RE, Jinek M, Wiedenheft B, Zhou K & Doudna JA (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355–1358.
13 Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF *et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467–477.
14 Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJ, van der Oost J, Doudna JA & Nogales E (2011) Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* **477**, 486–489.
15 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA & Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821.
16 Chylinski K, Le Rhun A & Charpentier E (2013) The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol* **10**, 726–737.
17 Sternberg SH, Redding S, Jinek M, Greene EC & Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67.
18 Swiech L, Heidenreich M, Banerjee A, Habib N, Li Y, Trombetta J, Sur M & Zhang F (2015) In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol* **33**, 102–106.
19 Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, Grompe M, Koteliansky V, Sharp PA, Jacks T & Anderson DG (2014) Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* **32**, 551–553.
20 Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K & Kuhn R (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol* **33**, 543–548.
21 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE & Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826.
22 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823.
23 Ratz M, Testa I, Hell SW & Jakobs S (2015) CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. *Sci Rep* **5**, 9592.
24 Zhu W, Lei R, Le Duff Y, Li J, Guo F, Wainberg MA & Liang C (2015) The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology* **12**, 22.
25 Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NG, van den Broek M, Daran-Lapujade P, Pronk JT, van Maris AJ & Daran JM (2015) CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. *FEMS Yeast Res* **15**. [Epub ahead of print].
26 Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S *et al.* (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997.
27 Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827–832.
28 Kiani S, Chavez A, Tuttle M, Hall RN, Chari R, Ter-Ovanesyan D, Qian J, Pruitt BW, Beal J, Vora S *et al.* (2015) Cas9 gRNA engineering for genome editing, activation and repression. *Nat Methods* **12**, 1051–1054.
29 Jinek M, East A, Cheng A, Lin S, Ma E & Doudna J (2013) RNA-programmed genome editing in human cells. *eLife* **2**, e00471.
30 Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK & Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822–826.

31 Kim S, Kim D, Cho SW, Kim J & Kim JS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* **24**, 1012–1019.
32 Shen B, Zhang J, Wu H, Wang J, Ma K, Li Z, Zhang X, Zhang P & Huang X (2013) Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res* **23**, 720–723.
33 Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ & Thomson JA (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proc Natl Acad Sci USA* **110**, 15644–15649.
34 Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JR *et al.* (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485.
35 Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu X, Makarova KS *et al.* (2015) In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, 186–191.
36 Chari R, Mali P, Moosburner M & Church GM (2015) Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat Methods* **12**, 823–826.
37 Wang T, Wei JJ, Sabatini DM & Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84.
38 Xu H, Xiao T, Chen CH, Li W, Meyer CA, Wu Q, Wu D, Cong L, Zhang F, Liu JS *et al.* (2015) Sequence determinants of improved CRISPR sgRNA design. *Genome Res* **25**, 1147–1157.
39 Moreno-Mateos MA, Vejnar CE, Beaudoin JD, Fernandez JP, Mis EK, Khokha MK & Giraldez AJ (2015) CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods* **12**, 982–988.
40 Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ & Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262–1267.
41 Lin S, Staahl BT, Alla RK & Doudna JA (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* **3**, e04766.
42 Wagner H (2001) Toll meets bacterial CpG-DNA. *Immunity* **14**, 499–502.
43 Gaj T, Guo J, Kato Y, Sirk SJ & Barbas CF 3rd (2012) Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat Methods* **9**, 805–807.
44 Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK & Kim H (2014) Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res* **24**, 1020–1027.
45 Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, Maeder ML, Joung JK, Chen ZY & Liu DR (2015) Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotechnol* **33**, 73–80.
46 Sun W, Ji W, Hall JM, Hu Q, Wang C, Beisel CL & Gu Z (2015) Self-Assembled DNA nanoclews for the efficient delivery of CRISPR-Cas9 for genome editing. *Angew Chem* **54**, 12029–12033.
47 Sung YH, Kim JM, Kim HT, Lee J, Jeon J, Jin Y, Choi JH, Ban YH, Ha SJ, Kim CH *et al.* (2014) Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res* **24**, 125–131.
48 Wang X, Wang Y, Wu X, Wang J, Wang Y, Qiu Z, Chang T, Huang H, Lin RJ & Yee JK (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* **33**, 175–178.
49 Holkers M, Maggio I, Henriques SF, Janssen JM, Cathomen T & Goncalves MA (2014) Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat Methods* **11**, 1051–1057.
50 Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C & Yusa K (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**, 267–273.
51 Mefferd AL, Kornepati AV, Bogerd HP, Kennedy EM & Cullen BR (2015) Expression of CRISPR/Cas single guide RNAs using small tRNA promoters. *RNA* **21**, 1683–1689.
52 Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA & Liu DR (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839–843.
53 Ma Y, Zhang L & Huang X (2014) Genome modification by CRISPR/Cas9. *FEBS J* **281**, 5186–5193.
54 Jiang W, Bikard D, Cox D, Zhang F & Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* **31**, 233–239.
55 Lin Y, Cradick TJ, Brown MT, Deshmukh H, Ranjan P, Sarode N, Wile BM, Vertino PM, Stewart FJ & Bao G (2014) CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res* **42**, 7473–7485.
56 Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, Wyvekens N, Khayter C, Iafrate AJ, Le LP *et al.* (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187–197.
57 Kim D, Bae S, Park J, Kim E, Kim S, Yu HR, Hwang J, Kim JI & Kim JS (2015) Digenome-seq: genome-

wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* **12**, 237–243. 1 p following 243.
58 Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X *et al.* (2014) Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods* **11**, 399–402.
59 Fu Y, Sander JD, Reyon D, Cascio VM & Joung JK (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**, 279–284.
60 Hara S, Tamano M, Yamashita S, Kato T, Saito T, Sakuma T, Yamamoto T, Inui M & Takada S (2015) Generation of mutant mice via the CRISPR/Cas9 system using FokI-dCas9. *Sci Rep* **5**, 11221.
61 Guilinger JP, Thompson DB & Liu DR (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol* **32**, 577–582.
62 Tsai SQ, Wyvekens N, Khayter C, Foden JA, Thapar V, Reyon D, Goodwin MJ, Aryee MJ & Joung JK (2014) Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol* **32**, 569–576.
63 Tomkinson AE, Howes TR & Wiest NE (2013) DNA ligases as therapeutic targets. *Transl Cancer Res* **2**.
64 Robert F, Barbeau M, Ethier S, Dostie J & Pelletier J (2015) Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med* **7**, 93.
65 Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR & Ploegh HL (2015) Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**, 538–542.
66 Basu S, Aryan A, Overcash JM, Samuel GH, Anderson MA, Dahlem TJ, Myles KM & Adelman ZN (2015) Silencing of end-joining repair for efficient site-specific gene insertion after TALEN/CRISPR mutagenesis in Aedes aegypti. *Proc Natl Acad Sci USA* **112**, 4038–4043.
67 Yu C, Liu Y, Ma T, Liu K, Xu S, Zhang Y, Liu H, La Russa M, Xie M, Ding S *et al.* (2015) Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**, 142–147.
68 Vartak SV & Raghavan SC (2015) Inhibition of nonhomologous end joining to increase the specificity of CRISPR/Cas9 genome editing. *FEBS J* **282**, 4289–4294.
69 Srivastava M, Nambiar M, Sharma S, Karki SS, Goldsmith G, Hegde M, Kumar S, Pandey M, Singh RK, Ray P *et al.* (2012) An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell* **151**, 1474–1487.
70 Deng W, Shi X, Tjian R, Lionnet T & Singer RH (2015) CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc Natl Acad Sci USA* **112**, 11870–11875.
71 Claussnitzer M, Dankel SN, Kim KH, Quon G, Meuleman W, Haugen C, Glunk V, Sousa IS, Beaudry JL, Puviindran V *et al.* (2015) FTO obesity variant circuitry and adipocyte browning in humans. *N Engl J Med* **373**, 895–907.
72 Parnas O, Jovanovic M, Eisenhaure TM, Herbst RH, Dixit A, Ye CJ, Przybylski D, Platt RJ, Tirosh I, Sanjana NE *et al.* (2015) A Genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675–686.
73 Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z, Lv J, Xie X, Chen Y, Li Y *et al.* (2015) CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein & Cell* **6**, 363–372.
74 Sander JD, Zaback P, Joung JK, Voytas DF & Dobbs D (2007) Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Res* **35**, W599–W605.
75 Naito Y, Hino K, Bono H & Ui-Tei K (2015) CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* **31**, 1120–1123.
76 Bae S, Park J & Kim JS (2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475.
77 Montague TG, Cruz JM, Gagnon JA, Church GM & Valen E (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res* **42**, W401–W407.
78 Heigwer F, Kerr G & Boutros M (2014) E-CRISP: fast CRISPR target site identification. *Nat Methods* **11**, 122–123.
79 Prykhozhij SV, Rajan V, Gaston D & Berman JN (2015) CRISPR multitargeter: a web tool to find common and unique CRISPR single guide RNA targets in a set of similar sequences. *PLoS One* **10**, e0119372.
80 Liu H, Wei Z, Dominguez A, Li Y, Wang X & Qi LS (2015) CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. *Bioinformatics* **31**, 3676–3678.