Genome modification by CRISPR/Cas9

Yuanwu Ma[1], Lianfeng Zhang[1], Xingxu Huang[2]*

[1]Key Laboratory of Human Disease Comparative Medicine, Ministry of Health, Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences, Beijing 100021, China.

[2]MOE Key Laboratory of Model Animal for Disease Study, Model Animal Research Center of Nanjing University, Nanjing Biomedical Research Institute, National Resource Center for Mutant Mice, Nanjing 210061, China.

*Correspondence: xingxuhuang@mail.nju.edu.cn (Xingxu Huang)

Article type : Review Article

Abstract

CRISPR/Cas9-mediated genome modification enables us to edit the genomes of a variety of organisms rapidly and efficiently. The advantages of the CRISPR/Cas9 system have made it an increasingly popular genetic engineering tool for biological and therapeutic applications. Moreover, CRISPR/Cas9 has been employed to recruit

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/febs.13110

functional domains that repress/activate gene expression or label specific genomic loci in living cells or organisms, in order to explore the developmental mechanisms, gene expression regulation and animal behavior. One major concern about this system is its specificity; although the CRISPR/Cas9-mediated off-target mutation has been broadly studied, more efforts are required to further improve the specificity of CRISPR/Cas9. We will also discuss the potential applications of CRISPR/Cas9.

Introduction

Precise modification of specific sites within a gene of interest is considered a standard approach to elucidate gene function, to create disease animal models and to improve desired characteristics of animals and plants. Targeted gene modification also provides the potential for therapeutic applications. In the past decades, strategies for precise genome modifications using embryonic stem (ES) cell-mediated modification by homologous recombination were limited to certain organisms. Recently, engineered nucleases including zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) have provided a much simpler and more economic way for gene targeted modification [1]. These engineered nucleases generate a DNA double-strand break (DSB) at the targeted genome locus. The break activates repair through error-prone non-homologous end joining (NHEJ) or homology-directed repair (HDR). In the absence of a template, NHEJ is activated, resulting in insertions and/or deletions (indels) which disrupt the target loci. In the presence of a donor

template with homology to the targeted locus, the HDR pathway operates allowing for precise mutations to be made.

The CRISPR/Cas system was first described as an adaptive immune system in bacteria and archaea, and has now been engineered as RNA guided endonucleases (RGENs) for genome editing. Typically, a type II CRISPR system functions by the CRISPR RNA (crRNA) interacting with a trans-activating crRNA (tracrRNA) to form a crRNA:tracrRNA duplex (this RNA duplex could be replaced with a single guide RNA, sgRNA), which directs CRISPR-associated protein (Cas9) from *Streptococcus pyogenes* to specific sites, thereby generating a DNA strand break [2]. Unlike ZFNs and TALENs, CRISPR/Cas9 system doesn't require the engineering of specific protein pairs for each target site, and introduces Cas9 to the target sequence based on RNA:DNA base-pairing rules. The simplicity of the CRISPR/Cas9 system has revolutionized genome engineering in a variety of cells and organisms.

In this review, we describe how the CRISPR/Cas9 system can be engineered for genome editing in different kinds of organisms. We will also discuss applications of CRISPR/Cas9 beyond genome editing, the improvement of the specificity of this system, and the challenges still remaining. Finally, we will highlight the bright future of this fascinating system in basic research and therapeutic applications.

To engineer CRISPR/Cas9 for genome editing

RNA-guide DNA cleavage systems protect bacteria and archaea against invading DNA

contaminants, serving as an adaptive immune system. The process is quite complicated. The invading foreign DNA can be recognized and inserted into a genome locus to form a CRISPR region; the captured foreign DNA sequences are termed protospacers. In type II CRISPR systems, the CRISPR locus is transcribed into a pre-CRISPR RNA (pre-crRNA) and processed to a matured crRNA with the assistance of trans-activating crRNA (tracrRNA). Interaction between crRNA and tracrRNA directs CRISPR-associated protein 9 (Cas9) of *Streptococcus pyogenes* to recognize the specific DNA sequence complementary with the protospacer. This RGEN target site is usually 20 bp in length and must be immediately adjacent to the NGG motif (or sometimes NAG, with much lower cleavage efficiency) [3] which are known as protospacer adjacent motifs (PAM). The programmable crRNA and fixed tracrRNA are fused to form a single guide RNA (sgRNA), which directs Cas9 to the desired site and catalyzes the cleavage of both DNA strands effectively.

The Cas9 protein contains the RuvC and HNH nuclease domains (Fig. 1). The HNH is a single domain while the RuvC domain consists of three subdomains. Single-particle electron microscopy reconstructions of *Streptococcus pyogenes* Cas9 (SpyCas9) showed a sgRNA-guided structural change to form a central channel for the RNA-DNA heteroduplex [2]. Later, the high-resolution structure of SpyCas9 in complex with guide RNA and target DNA showed a bilobed architecture including a target recognition (REC) lobe and a nuclease (NUC) lobe [4]. The nuclease lobe is composed of a HNH nuclease domain, a RuvC nuclease domain and a C-terminal region. The HNH and RuvC nuclease domains are responsible for the cleavage of the complementary and non-complementary DNA strands of the target sites (Fig. 1a) [5].

In general, Cas9 and sgRNA are the essential components for genome editing of the CRISPR/Cas9 system: sgRNA is responsible for the site targeting, and Cas9 contributes to the DNA cleavage at the target site. The PAM is required for the target site recognition [2]. Thus, any DNA sequence that contains the sequence of N_{20}-NGG could be recognized as target site. This system can be widely used in bacteria and many other kinds of organisms. Also, codon-optimized Cas9 with an appropriate nuclear localization signal (NLS) showed high activity in eukaryotes such as human cells, mice, and rats [6-8]. Since its first description two years ago, the sgRNA-guided Cas9 system has been applied to modify endogenous genes in a wide range of cells and organisms, including bacteria [9], yeast [10], plants [11-13], roundworm [14], silkworm [15], fruit fly [16], zebrafish [17], frog [18], rabbit [19], mouse [7, 20], rat [8, 21], pig [22], monkey [23], and different human cells [6, 24-26].

By providing multiple sgRNAs and programming the DNA-cleaving activity of Cas9, it has proved possible to simultaneously modify multiple genes in cells or animals. Four genes including *ApoE*, *B2m*, *Prf1*, and *Prkdc* have been disrupted simultaneously in rats at an efficiency of 24% [27], five genes including *B2m*, *Il2rg*, *Prf1*, *Prkdc*, and *Rag1* have been targeted in mice at an efficiency of 75% [28], five genes including *Tet1, 2, 3*, *Sry,* and *Uty* in mouse embryonic stem (ES) cells at 10% efficiency [20], and five genes including *Ddx*, *Egfp*, *Gol*, *Mitfa*, and *Tyr* in somatic cells of zebrafish also at high efficiency [29]. The multiple capacity of simultaneously disrupting multiple genes is especially useful for studying adjacent genes. Moreover, by exploiting the advantages of Cas9 used with multiple sgRNAs, large fragment deletion or inversion between sgRNA targeting sites have been successfully achieved [30]. When supplied with plasmid DNA or single-strand oligonucleotide as

templates, Cas9-assisted HDR operated and successfully generated site-specific mutation, *loxP* or reporter gene insertion in mice [31], *loxP* or reporter gene insertion in rats [32-33], and zebrafish [34].

The simplicity of the CRISPR/Cas9 system has also facilitated the generation of lentivirus-based sgRNA libraries covering almost all the mouse and human genes, which can be used for high-throughput functional screening [35-36]. Genome wide loss-of-function screens have already been applied to robust negative and positive selection screens in human cells [35, 37]. The sgRNA can be designed to target nearly any DNA sequence. Therefore, sgRNA libraries may also be used to study the function of noncoding genetic elements. Fusion of a non-functional or 'dead' Cas9 (dCas9, see below) with different effector domains has been used for the studies beyond loss-of-function phenotypes [38-39]. For example, dCas9 fused to an epigenetic modifier was applied to elucidate the methylation effect and certain chromatin states in defined conditions.

In addition, the CRISPR/Cas9 system has been used to correct the genetic disease in mice or in intestinal stem cell organoids of cystic fibrosis patients [40-41]. Other evidence showed that CRISPR/Cas9 corrected the *Fah* mutation in an adult mouse model of human hereditary tyrosinemia disease, and alleviated the symptoms of this disease [42].

The simplicity and high DNA-cleavage efficiency have made CRISPR/Cas9 an increasing popular genome editing tool. Nevertheless, potential users should pay attention to several points.

First, the delivery of sgRNA and Cas9. Both sgRNA and Cas9 are required for efficient target

site recognition and subsequent cleavage and to deliver both of them simultaneously is more difficult than either one alone. Different approaches such as electroporation, nucleofection, and lipofectamine have been used for the delivery of Cas9 and sgRNA expression plasmids into mammal cells. Lentiviral vectors have been used to construct large-scale sgRNA expressing libraries covering nearly all the human and mouse genes [35-36]. Meanwhile, *in vitro* transcribed Cas9 mRNA/plasmid DNA and sgRNA/plasmid DNA have been microinjected into the one-cell embryos of zebrafish [17], fruit flies [16], mice [7, 20], rats [8, 21, 32], pigs [22], and monkeys [23] to generate gene modified animals. The purified Cas9 protein complex together with sgRNAs also has been microinjected into one-cell embryos to generate knockout mice and zebrafish [43]. Time course experiment showed that Cas9 protein:sgRNA complex works earlier than Cas9 mRNA plus sgRNA [43]. Polyethylene glycol (PEG) and *Agrobacterium* have been used for Cas9 and sgRNA delivery into wheat, rice, sorghum, tobacco, and thale cress for genomic editing of these plants [11-13, 44]. The selection of a proper delivery system depends on the cells or organisms used.

Second, target site selection and sgRNA design. In theory, any DNA sequence that contains the sequence NGG (or NAG, to some lesser extent) is a potential target site. The RNA polymerase III-dependent U6 promoter and T7 promoter are most commonly used for the sgRNA expression. These promoters require a G or GG at the 5' end of the RNA to be transcribed. Therefore, a G or GG at the 5' end of the target site is required to start the transcription. A recent study reported a new vector using the U6 promoter to drive sgRNA expression without these limitations. In the vector, the U6 promoter drives the expression of multiple sgRNAs, each flanked by two Csy4 RNase cleavage sites from *Pseudomonas*

aeruginosa [45]. The expressed tandem multiple sgRNAs will be separated when Csy4 RNase is present [45]. This strategy allows any sequence containing the NGG at the 3' end to be used as the potential targeting sites [45]. In addition, a group reported that truncated sgRNAs with 17 or 18 nucleotides complementary with the target site also induce DNA cleavage efficiently. So sites in the form of $GN_{16\text{-}19}NGG$ or $GGN_{15\text{-}18}NGG$ can serve as the potential targeting sites. Such potential targeting sites theoretically will be found within the whole genome every 1 in 32 or 1 in 128 bp, respectively.

Third, the genotyping of CRISPR/Cas9-induced modifications. One pair of primers flanking the target site were designed to detect the modification in the target site using T7 endonucleases I assay (T7EN I) [7, 32]. Also, the same PCR products can be used for restriction fragment length polymorphism (RFLP) assay to detect the mutation [20, 31]. Detailed modifications were detected by sequencing analysis.

Finally, off-target effects. Besides, Cas9 induces mutations in its target and also off-target mutations too. Since the off-target events may cause unwanted modifications. It is very important to evaluate these in gene-modified cells. However, for animals, the side effects can be diluted by crossing with wild-type animals.

To improve the specificity

The CRISPR/Cas9-mediated off-target effects could be useful for bacteria and archaea to recognize and destroy those invading hyper-variable viral DNA or plasmid DNA. But for biological research or gene therapy, the off-target events will generate unwanted mutation

beyond the target site and result in side effects. Thus, increasing attention has been given to improving the specificity of CRISPR/Cas9 systems.

A number of studies on potential off-target effects showed that mismatches at the 5' end of the target site but not the "seed" (8-12 bp upstream of the PAM) region are generally better tolerated [2, 25]. However, mismatches less than three nucleotides and outside the 5' end also induce off-target events in human cells and rats [27, 32, 46]. Sometimes the off-target mutation occurs at a rate almost as high as that of the on-target cleavage [46]. Indeed, by analyzing the off-target events described in previously published papers, we found that mismatches of up to five nucleotides causes off-target mutation [46].

Deep sequencing assays were used to detect the off-target effects *in vivo* [45, 47-49]. Interestingly, most reported potential off-target sites were located at the noncoding regions in the host genomes [49]. The whole-exome sequencing assay showed no off-target event was found in CRISPR/Cas9-modified human cells [49]. Several CRISPR/Cas9 engineering online design and off-targeting searching tools have been developed, such as WTSI Genome Editing (WGE) (http://www.sanger.ac.uk/htgt/wge/), E-CRISP (http://www.e-crisp.org/E-CRISP), Genome engineering resources (http://www.genome-engineering.org/), RGENs tools (http://www.rgenome.net/), ZiFiT Targeter software (http://zifit.partners.org/ZiFiT/), GT-Scan (http://gt-scan.braembl.org.au/gt-scan/), CHOPCHOP (https://chopchop.rc.fas.harvard.edu) [1, 50-52]. However, a more comprehensive evaluation of the off-target effects mediated by CRISPR/Cas9 is still expected. For example, how many mismatches are tolerated for a given target site, and why do some of the potential off-target sites cause mutation while others do not? These differences may be caused by genomic/epigenomic context and/or chromatin

structure. We believe that a better understanding of the sgRNA target site screen is required for further improving the specificity of CRISPR/Cas9 system.

Great efforts have already been made to explore different strategies to reduce off-target effects.

(i) Select good target sites. Some guidelines for target site selection and sgRNA design are helpful for reducing the off-target effects. For example, it has been reported that high GC contents (up to 70%) in the target site could improve hybridization and tolerant more mismatches [53], while a high rate of off-target sites were observed in low GC content (less to 30%) [46, 54]. Mismatches that form DNA bulges at 5' end, 3' end or around 7~10 bp away from PAM [53] should be avoided, as well as potential sgRNA with bulges beyond the seed region [53].

(ii) Reduce the concentration of Cas9 and sgRNA. This strategy decreases the off-target effects, but may also affect on-target cleavages [55].

(iii) Use Cas9 mRNA/protein and sgRNA instead of Cas9 and sgRNA expression plasmid. Cas9 mRNA/protein and sgRNA work for shorter time, while plasmids keep expressing Cas9 and sgRNA which may increase the off-target effects and cause possible random integration of Cas9 and sgRNA into genome [1].

(iv) Truncate sgRNA at 3' end of tracrRNA and add two extra GG to the 5' end beyond the complementary region. Those two strategies decreased the ratio of off-target to on-target effects [54, 56].

(v) Use paired sgRNAs and Cas9 nickases. A Cas9 variant with D10A or H840A mutation in

Cas9 nucleases induces DSB between the paired sites by two sgRNAs (Fig. 1b) [57]. Targeting sites on opposite DNA strands separated by 4-100 bp would be recognized and cleaved by paired Cas9 nickases [50, 57]. Although several studies showed that Cas9 nickase with single sgRNA induce indel mutations at a very low level, the paired sgRNAs and Cas9 nickase do reduce the off-target effects in human cells and mice dramatically [58].

(vi) Truncate sgRNAs at the 5' end of prospacer regions. The tru-gRNAs have 17 or 18 bp of complementary sequence, which function effectively as full length sgRNA with improved specificity [47].

(vii) Use the dimeric CRISPR guide Fok I nuclease. Fusion of dCas9 to Fok I nuclease generated fCas9 (Fig. 1c) [45, 48]. DNA cleavage induced by fCas9 requires two sgRNAs targeted on opposite site of the DNA strands separated by 15 to 25 bp in the "PAM out" orientation. This modified fCas9 system showed comparable DNA modification efficiency to the Cas9 nickases, which is about two thirds of the efficiency of the wild-type Cas9, but dramatically increased the specificity [45, 48]. The studies of SpyCas9 structure also provide useful information for re-construction of Cas9 and sgRNA to improve the specificity [4-5].

Expanding the applications

As a versatile genetic engineering tool, CRISPR/Cas9 has been exploited beyond genome editing. The Cas9 variant containing D10A and H840A mutations is catalytically inactive or "dead" Cas9 (dCas9) (Fig. 1d). This dCas9 can be directed to the target site by sgRNA as effectively as wild type, but it can't function as nuclease for genome editing. This mutated

CRISPR/dCas9 has been applied for promoter targeting to repress gene expression in *Escherichia coli* and human cells, and also to recruit heterogeneous functional domains to a specific locus to repress/activate gene expression or label specific genomic loci in living cells or organisms (Fig. 1d). For example, dCas9 has been fused to the transcriptional activation domains such as VP64 or the p65 subunit to increase gene expression in human cells; and also fused to the Krüppel-associated box (KRAB) domain to decrease the gene expression [38-39, 59-61]. Such gene regulation can also be amplified in a synergistic way by using multiple sgRNAs. In a recent study, the fusion protein of TALE domains and LSD1 histone demethylase was used to regulate the enhancer-associated chromatin modifications [62]. Although no attempt has been made successfully, it will be very interesting to test whether dCas9 could be used for target epigenome editing by fusing with chromatin modification domains, such as DNA methylases, demethylases, histone acetylases, deaceylases and kinases.

An EGFP-tagged dCas9 has been used for imaging of the repetitive elements of telomeres and coding genes in living cells [63]. The labeling of specific genomic loci in living cells or organisms is a powerful strategy for the study of the spatiotemporal organization and dynamics of chromatin in regulating genome function [58]. The capability to regulate any endogenous gene will help us to pinpoint the factors responsible for cell differentiation and other cellular processes.

Conclusions and prospects

In summary, the simplicity and high efficiency of the CRISPR/Cas9 system allows affordable genome editing. In addition, the large sgRNA library will make both drug target identification and function screening more efficient. This RNA-guided genome editing tool also gives rise to the potential to change the genetic landscape of animals and plants around us to obtain the desired genotypes at will.

The application of this system has been expanded beyond genome editing, to areas such as gene expression regulation or specific chromatin labeling using fluorescent protein. Although great achievements have been made to improve the specificity and expand the application of this technology, there is still have plenty of room for improvement and extension.

The specificity of this system still needs improvement. A major concern for application of this system is off-target mutagenesis. In the past two years, many papers have reported off-target events of this system. Many efforts such as paired nickases, tru-gRNAs, and dimeric FokI-dCas9 nucleases were exploited to reduce off-target effects. Although these modifications have reduced the off-target events significantly, further improvement will be needed especially for the more precise modifications or therapeutic applications. Further optimization of this system needs unbiased strategies for more comprehensive evaluation of the off-target effects. Meanwhile, the mechanisms underlying target search remains unclear. Learning more about the mechanisms of the target search and cleavage will provide the basis for improvement of Cas9 and/or sgRNA to increase the specificity.

Another area for improvement is to shift the balance of HDR:NHEJ from NHEJ-mediated

indels towards HDR-mediated modifications. Although the site mutation or knockin event can be achieved at high rates in cells and animals by providing a single-stranded DNA or a double-strand DNA plasmid as template, how to reduce the NHEJ-mediated indels and improve the HDR-mediated precise modification are still the interesting issues. Co-localization of the template DNA with Cas9:sgRNA complex to the target site to enhance HDR or use the siRNA or inhibitor of NHEJ-mediated repair pathway to reduce the competing NHEJ should be very helpful.

Another, but not the last, important problem is to efficiently deliver CRISPR/Cas9 system into those cell types or tissues that are hard to transfect and/or infect. The development of safe gene delivery vehicles is necessary for the versatile use of CRISPR/Cas9 system.

Undoubtedly, the basic research will make its way into clinic practise. Further optimization and development of next-generation CRISPR/Cas9 tools for genome and epigenome editing is expected to satisfy the requirement of the therapeutic applications.

Acknowledgements

We thank Dr. Xin Lou from Model Animal Research of Nanjing University for careful reading and editing of the manuscript. We thank the entire Zhang Lab and Huang Lab for their support and advice. This work was supported by grants from the 973 program (2011CB944301, 2012BA139B02).

Reference

1 Kim H & Kim JS (2014) A guide to genome engineering with programmable nucleases. *Nat Rev Genet*, **15**, 321-334.

2 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.

3 Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L & Church GM (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*, **31**, 833-838.

4 Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S, Kaplan M, Iavarone AT, Charpentier E, Nogales E & Doudna JA (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, **343**, 1247997.

5 Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F & Nureki O (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, **156**, 935-949.

6 Jinek M, East A, Cheng A, Lin S, Ma E & Doudna J (2013) RNA-programmed genome editing in human cells. *Elife*, **2**, e00471.

7 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.

8 Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M, Li Y, Gao N, Wang L, Lu X & Zhao Y (2013) Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol*, **31**, 681-683.

9 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.

10 DiCarlo JE, Norville JE, Mali P, Rios X, Aach J & Church GM (2013) Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res*, **41**, 4336-4343.

11 Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM & Sheen J (2013) Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nat Biotechnol*, **31**, 688-691.

12 Nekrasov V, Staskawicz B, Weigel D, Jones JD & Kamoun S (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nat Biotechnol*, **31**, 691-693.

13 Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL & Gao C (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol*, **31**, 686-688.

14 Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM & Calarco JA (2013) Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat Methods*, **10**, 741-743.

15 Wang Y, Li Z, Xu J, Zeng B, Ling L, You L, Chen Y, Huang Y & Tan A (2013) The CRISPR/Cas System mediates efficient genome engineering in Bombyx mori. *Cell Res*, **23**,

1414-1416.

16 Bassett AR, Tibbit C, Ponting CP & Liu JL (2013) Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. *Cell Rep*, **4**, 220-228.

17 Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR & Joung JK (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*, **31**, 227-229.

18 Nakayama T, Fish MB, Fisher M, Oomen-Hajagos J, Thomsen GH & Grainger RM (2013) Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in Xenopus tropicalis. *Genesis*, **51**, 835-843.

19 Yang D, Xu J, Zhu T, Fan J, Lai L, Zhang J & Chen YE (2014) Effective gene targeting in rabbits using RNA-guided Cas9 nucleases. *J Mol Cell Biol*, **6**, 97-99.

20 Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F & Jaenisch R (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, **153**, 910-918.

21 Li W, Teng F, Li T & Zhou Q (2013) Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol*, **31**, 684-686.

22 Hai T, Teng F, Guo R, Li W & Zhou Q (2014) One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. *Cell Res*, **24**, 372-375.

23 Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, Kang Y, Zhao X, Si W, Li W, Xiang AP, Zhou J, Guo X, Bi Y, Si C, Hu B, Dong G, Wang H, Zhou Z, Li T, Tan T, Pu X, Wang F, Ji S, Zhou Q, Huang X, Ji W & Sha J (2014) Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos. *Cell*, **156**, 836-843.

24 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.

25 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA & Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819-823.

26 Cho SW, Kim S, Kim JM & Kim JS (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol*, **31**, 230-232.

27 Ma Y, Shen B, Zhang X, Lu Y, Chen W, Ma J, Huang X & Zhang L (2014) Heritable multiplex genetic engineering in rats using CRISPR/Cas9. *PLoS One*, **9**, e89413.

28 Zhou J, Shen B, Zhang W, Wang J, Yang J, Chen L, Zhang N, Zhu K, Xu J, Hu B, Leng Q & Huang X (2014) One-step generation of different immunodeficient mice with multiple gene modifications by CRISPR/Cas9 mediated genome engineering. *Int J Biochem Cell Biol*, **46**, 49-55.

29 Jao LE, Wente SR & Chen W (2013) Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A*, **110**, 13904-13909.

30 Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, Zu Y, Li W, Huang P, Tong X, Zhu Z, Lin S & Zhang B (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res*, **41**, e141.

31 Yang H, Wang H, Shivalila CS, Cheng AW, Shi L & Jaenisch R (2013) One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated

genome engineering. *Cell*, **154**, 1370-1379.
32 Ma Y, Zhang X, Shen B, Lu Y, Chen W, Ma J, Bai L, Huang X & Zhang L (2014) Generating rats with conditional alleles using CRISPR/Cas9. *Cell Res*, **24**, 122-125.
33 Ma Y, Ma J, Zhang X, Chen W, Yu L, Lu Y, Bai L, Shen B, Huang X & Zhang L (2014) Generation of eGFP and Cre knockin rats by CRISPR/Cas9. *FEBS J*, **281**, 3779-3790.
34 Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X, Xiong JW & Xi JJ (2013) Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res*, **23**, 465-472.
35 Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG & Zhang F (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, **343**, 84-87.
36 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.
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 Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS & Qi LS (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, **154**, 442-451.
39 Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM & Zhang F (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, **500**, 472-476.
40 Wu Y, Liang D, Wang Y, Bai M, Tang W, Bao S, Yan Z, Li D & Li J (2013) Correction of a Genetic Disease in Mouse via Use of CRISPR-Cas9. *Cell Stem Cell*, **13**, 659-662.
41 Schwank G, Koo BK, Sasselli V, Dekkers JF, Heo I, Demircan T, Sasaki N, Boymans S, Cuppen E, van der Ent CK, Nieuwenhuis EE, Beekman JM & Clevers H (2013) Functional Repair of CFTR by CRISPR/Cas9 in Intestinal Stem Cell Organoids of Cystic Fibrosis Patients. *Cell Stem Cell*, **13**, 653-658.
42 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.
43 Sung YH, Kim JM, Kim HT, Lee J, Jeon J, Jin Y, Choi JH, Ban YH, Ha SJ, Kim CH, Lee HW & Kim JS (2014) Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res*, **24**, 125-131.
44 Jiang W, Zhou H, Bi H, Fromm M, Yang B & Weeks DP (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res*, **41**, e188.
45 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.
46 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.
47 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.

48 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.

49 Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S & Kim JS (2013) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*, **24**, 132-141.

50 Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, Wang L, Hodgkins A, Iyer V, Huang X & Skarnes WC (2014) Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods*, **11**, 399-402.

51 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**, W401-407.

52 O'Brien A & Bailey TL (2014) GT-Scan: Identifying unique genomic targets. *Bioinformatics*, **30**, 2673-2675.

53 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.

54 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.

55 Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G & Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*, **31**, 827-832.

56 Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S & Kim JS (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*, **24**, 132-141.

57 Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y & Zhang F (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, **154**, 1380-1389.

58 Sander JD & Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*, **32**, 347-355.

59 Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH & Joung JK (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat Methods*, **10**, 977-979.

60 Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW, Guilak F, Crawford GE, Reddy TE & Gersbach CA (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods*, **10**, 973-976.

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

62 Mendenhall EM, Williamson KE, Reyon D, Zou JY, Ram O, Joung JK & Bernstein BE (2013) Locus-specific editing of histone modifications at endogenous enhancers. *Nat*

Biotechnol, **31**, 1133-1136.

63 Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS & Huang B (2013) Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*, **155**, 1479-1491.

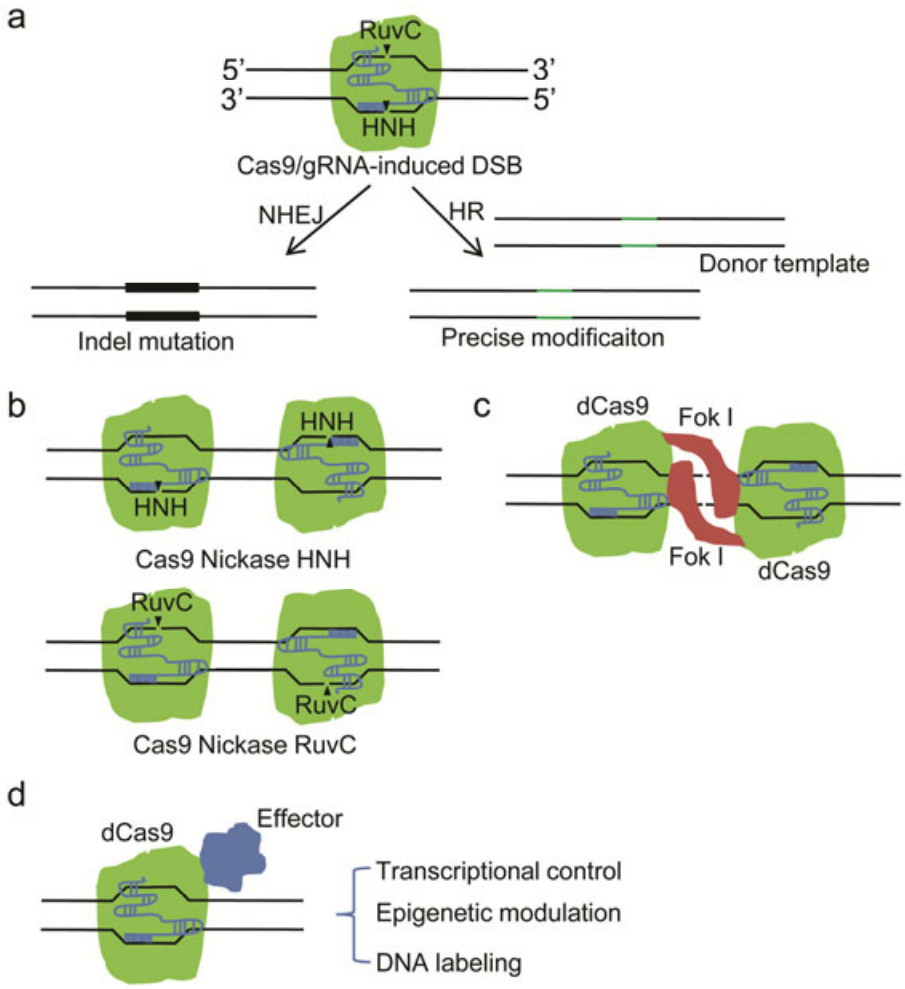

Figure 1 CRISPR/Cas9-mediated genome editing

a. Cas9/sgRNA induced double strand breaks (DSBs) can be repaired by either non-homologous end joining (NHEJ) or homology-directed repair (HDR) pathways. The Cas9 contains RuvC and HNH nuclease domains, each of which is responsible for one strand

DNA cleavage. b. Paired nickases were used to improve the specificity in the genome editing. Cas9 nickase (HNH) cleaves only the strand DNA (complementary strands of the target DNA) recognized by the sgRNA. Cas9 nickase (RuvC) cleaves the strand DNA (non-complementary strands of the target DNA) not interacting with the sgRNA. c. dCas9 ('dead' Cas9, both HNH and RuvC nuclease domain are inactivated by mutation) is fused with Fok I nuclease to improve the specificity of genome editing. d. dCas9 fused with an effector domain, such as DNA methylases, demethylases, histone acetylases, deaceylases, and kinases to modify the specific chromatin modification for desired effects.