

REVIEW

Advances in genetic engineering of domestic animals

Shaohua WANG¹, Kun ZHANG (✉)¹, Yunping DAI (✉)²

¹ Laboratory of Mammalian Molecular Embryology, College of Animal Sciences, Zhejiang University, Hangzhou 310058, China

² State Key Laboratory for Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing 100193, China

Abstract Global population will increase to over nine billion by 2050 with the doubling in demand for meat and milk. To overcome this challenge, it is necessary to breed highly efficient and productive livestock. Furthermore, livestock are also excellent models for human diseases and ideal bioreactors to produce pharmaceutical proteins. Thus, genetic engineering of domestic animals presents a critical and valuable tool to address these agricultural and biomedical applications. Overall, genetic engineering has evolved through three stages in history: transgenesis, gene targeting, and gene editing. Since the birth of the first transgenic pig, genetic engineering in livestock has been advancing slowly due to inherent technical limitations. A major breakthrough has been the advent of somatic cell nuclear transfer, which, for the first time, provided the technical ability to produce site-specific genome-modified domestic animals. However, the low efficiency of gene targeting events in somatic cells prohibits its wide use in agricultural and biomedical applications. Recently, rapid progress in tools and methods of genome engineering has been made, allowing genetic editing from mutation of a single base pair to the deletion of entire chromosomes. Here, we review the major advances of genetic engineering in domestic animals with emphasis placed on the introduction of latest designer nucleases.

Keywords CRISPR, TALEN, ZFN, gene editing, livestock, pig, cattle

Division of United Nations). In particular, given the rapid economic development and urbanization in developing countries, including China, food composition is changing with the demand for animal protein increasing dramatically. Specifically, the FAO (the Food and Agricultural Organization of the United Nations) predicts that global demand for meat and milk will double by 2050. Thus, highly efficient and productive livestock are needed to meet this increased demand. Nonetheless, traditional breeding of livestock is limited by the length of the process, slow progress of genetic improvements and inability to isolate desired traits from undesirable traits^[1].

To overcome these limitations, it is necessary to take advantage of advances in biotechnology, such as gene engineering (Fig. 1). For instance, expression of lysostaphin specifically in the bovine mammary gland makes the cow highly resistant to infection by *Staphylococcus aureus*. Extensive efforts in our and other laboratories have been made to improve meat production^[4,5], optimize milk composition with improved nutritional value^[6–8].

Domestic animals also present excellent models in biomedical research (Fig. 1). For example, pigs are widely used as models for human diseases, including diabetes, Alzheimer's disease, cystic fibrosis and cardiovascular diseases^[9]. Furthermore, pigs are proposed as potential sources of organs to address the shortage of donated organs for transplantation in humans^[10]. Cows have also been genetically modified as bioreactors to produce pharmaceutical proteins^[6]. Currently, the most popular method to produce pharmaceutical proteins is to use a mammalian cell culture system, which is expensive and limited by its capacity, whereas the use of bovine mammary gland for the expression of human genes promises to be a cost-effective method to produce valuable pharmaceutical proteins^[11].

Advances in animal genetics, reproduction and nutrition have resulted in a significant increase in the productivity of modern livestock. However, further development of these fields requires a better understanding of the identity and function of critical genes affecting important economic traits. Meanwhile, with the completion of whole genome sequencing for various mammals, functional annotation of their genomes remains elusive, especially in livestock due

1 Introduction

There is an old Chinese saying: food is the paramount necessity of the people (in other words, hunger breeds discontentment) indicating the importance of food security. Worldwide, the population is still growing with an estimate of nine billion by 2050, which means a requirement for a 50% increase in food supply (Data from the Population

Received January 8, 2016; accepted February 8, 2016

Correspondences: kzhang@zju.edu.cn, daiyunping@sina.com

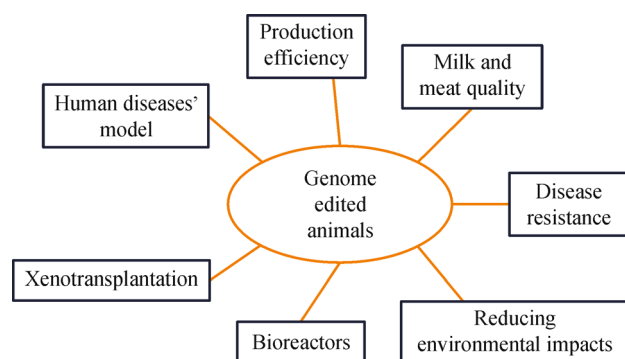


Fig. 1 Applications of genome editing in livestock

to limitations in “loss of function” tools. Indeed, mouse genetics has been greatly enhanced by the use of site-specific genome mutagenesis in embryonic stem cells (ESCs) followed by production of knockout animals. However, reliable ESCs have yet to be established for livestock.

To fulfill the demand for agricultural and biomedical applications as well as fundamental research in animal genetics, extensive efforts have been made to advance the genome modification technologies in livestock (Fig. 2). From the middle 1980s to the early 1990s, pronuclear DNA microinjection was the most popular approach to generate genome modification, albeit in a random-integration manner^[12,13]. Thanks to the birth of the cloned sheep, Dolly, gene targeted animals could be produced by performing site-specific genome modification in somatic

cells followed by nuclear transfer^[14]. This strategy is technically challenging, however, preventing its wide use in functional characterization of livestock genomes. Recently, application of designer nucleases, particularly the clustered regularly interspaced short palindromic repeats-associated protein 9 (CRISPR-Cas9) system, in site-specific genetic engineering provides a revolutionary tool for animal biotechnology and functional genomics in domestic animals^[15,16].

In this review, we will summarize the major advances in genome engineering of livestock over the last three decades. Emphasis is placed on the recent introduction and applications of designer nucleases, including zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), CRISPR-Cas9 in livestock.

2 History of genetic engineering in animals

2.1 Phase one: random-integrated transgenesis

In the 1970s, the earliest trials of genomic modification in mammals were conducted in mouse after the first application of genetic engineering in bacteria^[17]. The major advances in the field of molecular biology (e.g., recombinant DNA) and feasibility of culturing and manipulating mouse embryos made it possible to perform mouse transgenesis during that period. Several groups independently generated transgenic mice during the late 1970s and the early 1980s^[13,18–20]. Among them, the first

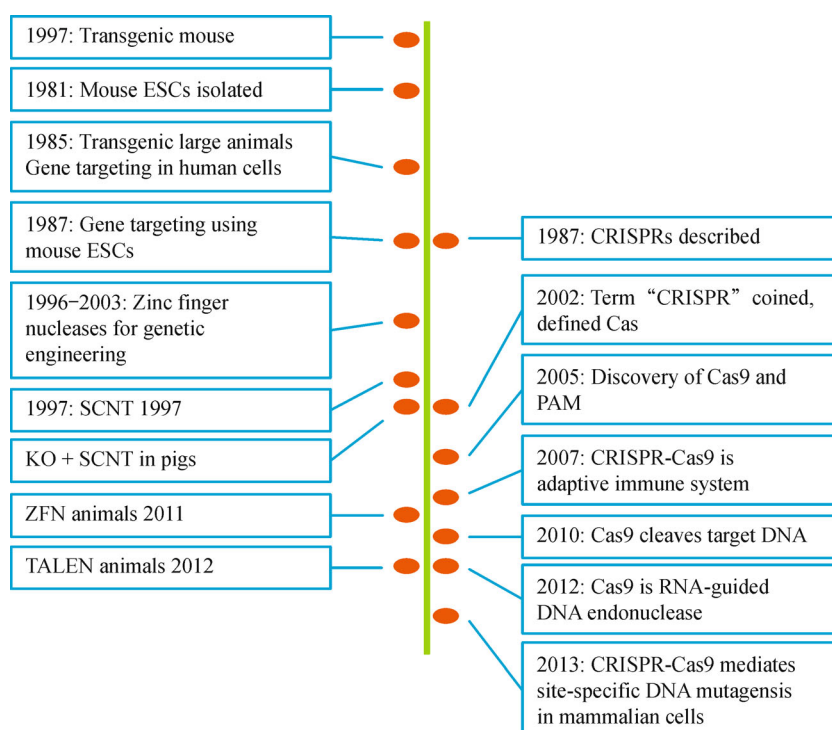


Fig. 2 Milestone events in the history of genetic engineering and CRISPR technology

transgenic mouse that resulted in changes in phenotype expressed rat growth hormone (GH)^[13]. Pronucleus DNA microinjection was adopted as the most popular method to deliver foreign gene to the host genome^[18]. The first report of transgenesis in livestock animals, involved microinjection of the gene encoding human growth hormone into the pronuclei or nuclei of eggs from super-ovulated rabbits, sheep and pigs^[12]. Integration of the gene into the host genome was demonstrated in all three species. Six years later, the first transgenic cattle were generated via similar approach^[21]. This technique had low efficiency, however, with only 1% of injected zygotes producing transgenic animals^[22].

Meanwhile, several alternative delivery methods were explored. Retrovirus infection: genetically modified animals could be produced by infecting early embryos with a replication-deficient virus carrying the transgene. Indeed, employing this approach in pigs dramatically enhanced the proportion of transgenic offspring obtained^[23,24]. Sperm-mediated DNA transfer: sperms are capable of integrating foreign DNA during fertilization, thus allowing highly efficient production of transgenic animals^[25]. These two methods are technically easy and relatively cheap to deliver a foreign gene into the host genome. Nonetheless, disadvantages include the difficulty in separating embryos for transgene integration before embryo transfer, the lack of specificity of expression because of random integration of exogenous DNA, and the inability to perform other types of gene editing (e.g., deletion) rather than transgene insertion.

The integration of exogenous genes into the host genome is random due to its reliance on double-stranded DNA break repair machinery. As a result, some embryos failed to develop when the foreign gene was inserted into sites that interfere with the expression of critical genes. Moreover, when transgene integration occurred within heterochromatin, expression of the transgene could be suppressed and modulated by those adjacent regulatory elements. With approximately 15% of microinjected livestock embryos resulting in live offspring, and a success rate of less than 1% of those embryos resulting in transgenic offspring, embryo survival and transgene integration have been major hurdles in the advancement of transgenic animal production^[26]. These inherent problems limit its extensive use as an important tool in addressing fundamental questions in life science as well as agricultural and biomedical applications.

2.2 Phase two: homologous recombination-directed gene targeting

The principle of homologous recombination (HR) was first documented in bacteria over 50 years ago. In the 1970s, similar machinery was also identified in eukaryotes^[27]. It was subsequently shown that HR-mediated recombination events occur in mammalian cells^[28]. Taking advantage of

this endogenous mechanism, it was used to deliver a plasmid DNA sequence into the genome of a human cell^[29]. Another milestone in the history of mammalian genetic engineering, was the successful culture and manipulation of mouse ESCs *in vitro* during the early 1980s^[30]. Thus, ESCs were soon adapted to create site-specific genetically engineered animals by injecting the genetically targeted cells into blastocysts^[31,32].

To date, ESCs have not been established for livestock, despite extensive efforts having been made, prohibiting its large-scale applications in gene targeting of livestock. Fortunately, a revolutionary tool, somatic cell nuclear transfer (SCNT), was successfully implemented with the birth of Dolly in mid-1990s^[33]. For the first time, SCNT presents a methodological basis for performing site-specific/targeted genetic modifications in livestock animals^[14,34]. Nonetheless, in comparison with mouse ESCs, somatic cells have limited life span in cell culture and low HR activity (one desired recombinant event per 10⁶ cells)^[35,36], which present an enormous technical challenge to conduct targeted genetic modifications in somatic cells. Despite this low efficiency, Prather and his colleagues obtained the first pig carrying site-specific genetic modifications using this approach^[14]. Shortly afterwards, the first cattle with site-specific genetic modifications were produced^[37].

A variety of strategies have also been developed to improve the efficiency of gene targeting via HR. These methods include positive and negative selection^[38], and gene trapping^[14]. Furthermore, the flanking homologous arms were extended via BAC vectors to improve HR efficiency. For instance, the usage of modified BAC vectors in targeting CFTR and DMD^[39,40] porcine kidney cells led to increased targeting efficiency (above 1%).

2.3 Phase three: designer nuclease-mediated gene editing

A breakthrough in the efforts to identify approaches to enhance HR was the discovery that the efficiency of HR-mediated gene targeting in mouse ESCs can be improved significantly by the targeted introduction of double strand break (DSBs) created by a rare-cutting endonuclease *I-SceI*^[41–43]. Eventually, experiments with zinc finger proteins demonstrated its potential as a designer nuclease to generate a DSB at a site of interest^[44]. So far, three major designer nucleases have been identified, namely, ZFNs, TALENs and RNA-guided endonucleases derived from the bacterial CRISPR-Cas system.

All these nucleases share similar mechanisms in mediating site-specific gene editing. They produce site-specific DNA DSBs, which elicit the endogenous DNA repair system. Generally, there are two major repair mechanisms: non-homologous end-joining (NHEJ) and homology-directed repair (HDR), both of which can be used for targeted editing as defined below^[45] (Fig. 3).

Gene disruption: In absence of homology arms, the

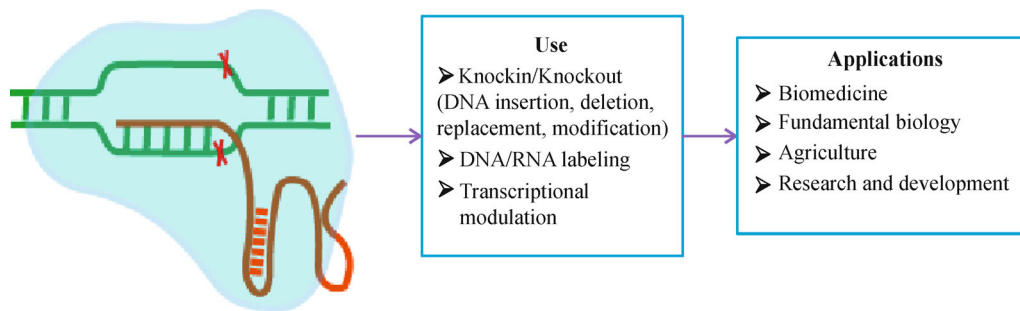


Fig. 3 Various applications of CRISPR-Cas9 technology

error-prone NHEJ pathway can introduce short insertions and deletions (indels) in the target site, eventually resulting in frame-shift mutations and gene disruption.

Gene insertion: Editing can be achieved by delivering designer nucleases along with a targeting vector carrying the genetic segment to be inserted and flanked by homology arms.

Single nucleotide mutagenesis: Single nucleotide/point mutations can be corrected or introduced in the genome via delivery of both designer nucleases and targeting vectors or single-strand oligonucleotides.

Rearrangements of chromosomes: DSBs stimulated by designer nucleases can also trigger chromosomal rearrangements, such as deletions, duplications and inversions.

2.3.1 Zinc finger nucleases

ZFNs have two components: nonspecific DNA cleavage domains, derivatives of *FokI* endonuclease and site-specific zinc finger proteins, which are naturally occurring and the most abundant DNA binding domains in humans. Each zinc finger domain can recognize 3 bp of DNA. Two decades ago, the first crystal structure of zinc finger proteins was disclosed, showing simplicity of the interaction with DNA^[46]. Since then, extensive efforts have been conducted to explore methods to engineer these proteins with sequence-specificity. Eventually, zinc finger proteins can be assembled with user-defined specificity, although this is time-consuming and expensive. Up to now, ZFN-induced gene editing has been achieved in a variety of organisms, including *Drosophila*^[47], *Caenorhabditis elegans*^[48], zebrafish^[49], rats^[50], cattle^[5,51] and pigs^[52,53].

In 2011, we successfully obtained β -lactoglobulin in knockout cattle via a ZFN-induced gene-targeting strategy with high efficiency^[51]. β -lactoglobulin belongs to a major whey protein in bovine milk and has been found as a major milk allergen. Eight cattle were produced from the bi-allelically targeted donor cells. However, only one calf survived because of the high-frequency of calf loss resulting from animal cloning. The high rate of induced bi-allelic mutations greatly reduces the time to produce

null mutant livestock animals and was a great stride for animal genome modification.

In our laboratory, ZFNs were also assembled to disrupt the bovine *MSTN* gene that encodes myostatin protein^[5]. Myostatin belongs to TGF- β superfamily and inhibits the growth of muscle cells. Using a ZFN-induced gene targeting strategy, the frequency of mutation is around 20% and 8% bi-allelic mutation efficiency is achieved. By performing SCNT, we successfully obtained cloned cattle from the *MSTN*-targeted fibroblast cells derived from Chinese yellow cattle^[5]. As expected, these cattle exhibit a double muscling phenotype as early as one month old and muscle fiber hypertrophy is also noted.

Although the application of ZFNs opened a new era of gene targeting, it has not been widely used due to the difficulty in engineering active ZFNs and validating them for target sites of interest. Another disadvantage is its inability to target DNA sites at a resolution of a single nucleotide (e.g., SNPs and active sites of enzymes) due to its characteristic of context-dependent specificity in recognizing DNA.

2.3.2 Transcription activator-like effector nucleases

Transcription activator-like effector (TALE) proteins exist naturally only in *Xanthomonas* bacteria^[54,55]. TALENs, like ZFNs, are artificial fusion proteins that contain a sequence-specific DNA binding motif (TALE domain) and nonspecific *FokI* endonuclease. It has become a valuable tool in genome engineering since the TALE-DNA binding code was determined in 2009^[55,56]. TALEs, like ZFNs, consists of an array of module repeats, each of them having 34 amino acids versus 30 for zinc fingers. Each TALE module is able to recognize one nucleotide of DNA. The use of TALEN in genome engineering was first reported in 2010^[57] and by 2012, it was used to produce genetically modified livestock^[58].

As one of the first studies in using TALENs to modify the genomes of livestock, Carson et al. demonstrated that TALENs are very efficient in making targeted DNA modifications in the bovine genome in both fibroblast cells

and preimplantation embryos^[58]. When fibroblast cells were co-transfected with two TALEN pairs targeting the same chromosome in pigs, large deletions and inversions of chromosomal segments were found to occur^[58]. Also, MSTN-targeted cattle can be made simply through microinjection of TALEN mRNAs to bovine zygotes, greatly reducing the time needed to produce homologous gene-targeted livestock.

Compared with ZFNs, TALENs are easy to design and relatively cheap to assemble commercially (Table 1). However, the difficulty in design, synthesis and validation of TALEN proteins remains a major challenge that limits its widespread application in genome engineering.

2.3.3 Clustered regularly interspaced palindromic repeats

First described in 1987, CRISPRs consist of an array of short repetitive sequences (29 nt direct repeats) interspaced by short sequences (32 nt) in the *Escherichia coli* genome^[60]. It is noteworthy that similar repeat elements have also been found in numerous other bacterial and archaeal species. These sequences were later classified as unique repeat elements existing in more than 40% of sequenced bacteria and the majority of archaea^[61]. The term CRISPR was coined later and Cas genes were identified located adjacent to each CRISPR locus^[62]. However, the functional role of CRISPR and Cas9 in prokaryotes remained unknown.

A breakthrough observation was published in 2005, indicating that the spacer sequences separating the direct repeats within CRISPRs originated from plasmid and phage-associated sources^[63–65]. Previous studies also demonstrated that (1) CRISPR loci are transcribed^[66], (2) Cas proteins carry putative nuclease and helicase domains^[62], and (3) viruses cannot infect archaeal cells that contain spacer sequences matched snippets of their own DNA^[63]. Based on this evidence, it was proposed that CRISPR-Cas is an adaptive immunity system that defends prokaryote cells against phage infection^[63,64]. This hypothesis was confirmed in 2007 by the infection of the lactic acid bacterium *Streptococcus thermophilus* with

lytic phages^[67]. In 2008, mature CRISPR RNAs (crRNAs) were shown to serve as guides in a complex with Cas proteins to interfere with virus infection in *E. coli*^[68].

It was originally presumed that spacers within CRISPR interfere with viral gene expression in a similar manner to RNAi. In 2008, it was demonstrated in *Staphylococcus epidermidis* that the target of Cas enzyme activity is DNA, not RNA, ruling out the possibility of an RNAi-like mechanism for the CRISPR-Cas system^[69]. Two years later, Moineau and colleagues discovered that the CRISPR-Cas9 system creates DSBs in the exact position of the target DNA 3 base pairs upstream of the protospacer adjacent motif (PAM) sequence, highlighting the critical role of the PAM sequence^[70]. Moreover, Cas9 is the only protein within the Cas gene cluster that is required for targeted DNA cleaving in the CRISPR-Cas9 system^[70]. Clarifying the final uncertainty about the mechanism of CRISPR-Cas9-mediated immunity system, Charpentier and colleagues observed a second non-coding RNA called trans-activating CRISPR RNA (tracrRNA) distinct from crRNA as determined by small RNA sequencing of *Streptococcus pyogenes*^[71]. In addition, they found that tracrRNA hybridizes with crRNA to form a duplex, which guides Cas9 to its targets^[71].

3 Applications of CRISPR-Cas9-induced genome editing

In 2013, the CRISPR-Cas9 system was adopted for genome editing in mammalian cells. Zhang and colleagues constructed two different Cas9 orthologs and revealed targeted DNA cleavage in both mouse and human cells^[15]. It was also been shown that this system could be harnessed to target multiple DNA loci and mediate HDR^[15]. At the same time, another laboratory from Harvard University reported similar findings^[16]. Since then, CRISPR-Cas9 has been widely employed as a preferred tool to perform genome editing in a variety of cells and organisms, including mice^[72], rats^[73], pigs^[74] and monkeys^[75], because of its simplicity in design and synthesis (Fig. 4).

Table 1 Comparison of different designer nucleases

ITEM	ZFNs	TALENs	RNA-guided ¹ nucleases
Genome targeting mechanism	Zinc-finger proteins	TALENs	sgRNA
Nuclease	<i>FokI</i>	<i>FokI</i>	Cas9 (wildtype or variants)
Mutationrate ²	10%	20%	20%
Requirement of target design	G-rich	T must be the first and A the last	End with NGG/NAG (PAM)
Off-target events	Frequent	Infrequent	Variable
Side effects	Variable	Low	Low
Size	1 kb	3 kb	4.2 kb (spCas9) + 0.1 kb (sgRNA)
Commercial price ^[59] (\$/target)	4000–7000	3600–5000	500

Note: 1, Includes wildtype Cas9 and its variants; 2, defined as the proportion of NHEJ-mediated indels created at the target DNA sites.

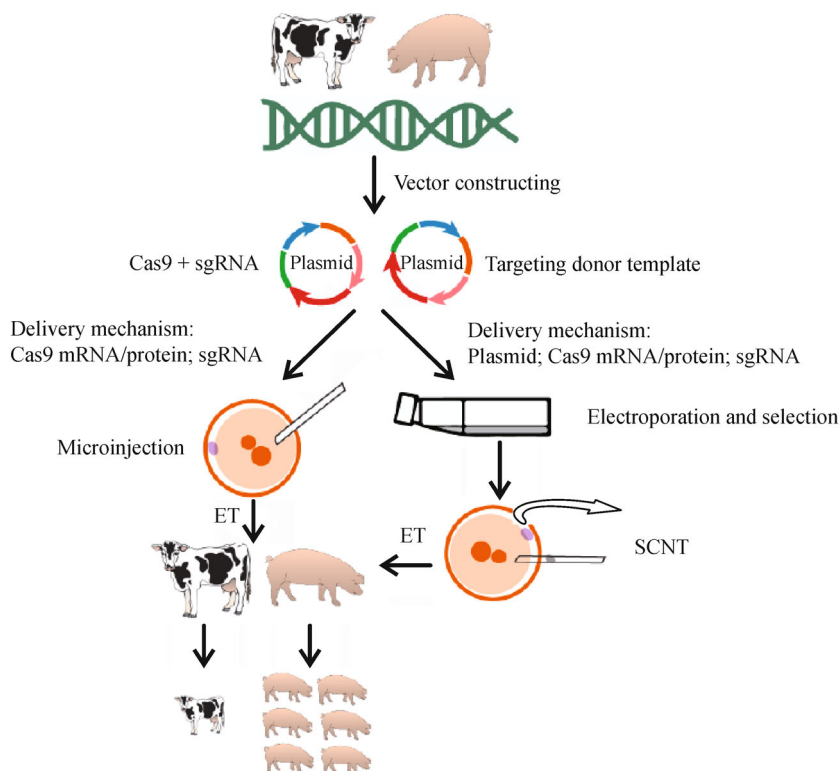


Fig. 4 Overview of producing genome modified animals with CRISPR-Cas9 technology

Typically, there are two essential components of a customized CRISPR-Cas9 system: non-specific Cas9 enzyme and sgRNA (single guide RNA), a designed RNA that mimics the crRNA-tracrRNA hybrid and guides Cas9 to target sites. A prerequisite for selecting target sites is the presence of PAM immediately downstream of the target site because of the essential role of PAM in mediating DNA cleavage activity of Cas9. Since the initial report of the use of CRISPR-Cas9 in genome editing, extensive efforts have been made to optimize the system. For example, optimizing the Cas9 codon and engineering of Cas9 by adding a nuclear localization signal, leading to improved gene editing activity in mammalian cells or organisms, e.g., human cells, mice, and rats.

One distinguishing feature of the CRISPR-Cas9 system is the capacity to simultaneously cleave multiple genes. This can be achieved by co-expression of an array of CRISPR carrying multiple spacers or multiple sgRNAs that target distinct sequences. With this strategy, efficient gene editing at multiple DNA sites has been demonstrated in mammalian cells and early embryos^[16,72]. Furthermore, deletion and inversion of large chromosomal segment can be completed by this approach.

Cas9 protein has also been engineered to destroy its DNA cleavage activity while maintaining its DNA binding activity. This modified type of Cas9, also termed dead Cas9 (dCas9), are being explored for multiple purposes in molecular biology as a RNA-guided target DNA binding

protein. For instance, dCas9 can be used to modulate transcription^[76–78]. Indeed, transcription initiation or elongation can be inhibited when dCas9 is guided to bind the promoter region or open reading frame in human cells^[76,79]. Fusion of dCas9 with the transcriptional repression domain of KRAB will lead to significantly reduced transcription in either yeast or human cells^[76,80]. Moreover, tagging dCas9 with fluorescent proteins can be used to visualize specific DNA loci in live cells^[78]. For example, Chen and coworkers successfully observed the spatiotemporal dynamics of DNA sequences of interest in live cells by fusing dCas9 with Egfp^[81].

Using the CRISPR-Cas9 system, knockout pigs have been successfully produced by either SCNT^[82] or directly by microinjection of zygotes^[74,82]. Zhou and his colleagues first demonstrated the feasibility of efficiently producing bi-allelic knockout pigs simply through injection of Cas9 mRNA and sgRNA into zygote cytoplasm^[74]. Sixteen piglets were generated, including six bi-allelic mutants and five mono-allelic mutants, indicating a high efficiency of gene targeting. Meanwhile, they also addressed low toxicity of the genome manipulation to the porcine embryonic development as evidenced by the high birth rate and survival rates^[74].

The generation of CD163 and CD1D knockout pigs using the CRISPR-Cas9 system has also been reported^[82]. Astonishingly, 100% efficiency of Cas9-induced targeted gene modification was achieved in injected zygotes. As in

mouse and human cells, it was also feasible to disrupt multiple genes in parallel by delivering CRISPRs targeting multiple genes into zygotes produced *in vitro*^[82].

By comparison with ZFNs and TALENs, the CRISPR-Cas9 system holds similar specificity and, more importantly, it is relatively easy to design and synthesize with minimal cost^[83] (Table 1). The advent of CRISPR-Cas9 tools has indeed revolutionized the creation of targeted gene modified animals. Previously, it typically took years to obtain pigs carrying bi-allelic mutants but now it takes only 6 months.

4 Specificity of CRISPR-Cas9 technology

Off-target effects are a major concern for the application of the CRISPR-Cas9 system since the off-target events may lead to unwanted genome modifications. To overcome this challenge, a variety of methods (detailed below) have been developed to improve the targeting efficiency of the CRISPR-Cas9 system.

4.1 Selection of target sites

Multiple factors related to selecting target sites and designing sgRNAs have been found to affect the frequency of off-target events. For instance, high GC content of the target site is recommended to enhance hybridization and accommodate more mismatches because low GC contents were associated with high frequency of off-target events^[84].

4.2 Selecting delivery content (DNA/RNA/protein)

This is helpful for reducing off-target effects by delivering Cas9 mRNA/protein and sgRNA rather than Cas9 and sgRNA expression plasmids. mRNA/protein only work temporarily, however, plasmids will stay for a relatively long time to express Cas9 and sgRNA, which increases the frequency of off-target events and leads to more random integration of CRISPR-Cas9 into the host genome^[85].

4.3 Using double-nicking approach

Cas9 enzymes cut DNA by using their HNH and RuvC nuclease domains, each one cleaving a strand of DNA to produce a blunt-ended DSB. Cas9 can be engineered into a nickase that only generates a single-stranded break (SSB) by introducing mutations (D10A or H840A) to inactivate one of those two nuclease domains. DNA SSBs are typically repaired through the base excision repair pathway, which exhibit high fidelity. Thus, paired sgRNAs and Cas9 variants with D10A or H840A mutation can be harnessed to induce DSBs and produce efficient indel formation. This strategy would greatly enhance the specificity compare to wildtype Cas9 due to the fact that

those off-target nick sites can be precisely repaired through the BER pathway^[86].

4.4 Truncated sgRNAs

It has been reported that truncated sgRNA with less than 20 nt complementary sequences significantly reduces the chance of undesired mutagenesis at off-target sites^[87]. Moreover, if this strategy was used in conjunction with a double-nicking approach, the off-target effects will be further reduced.

5 Conclusions

Before the advent of designer nucleases as tools for genome editing, the dissection of gene function was restricted in higher organisms by a lack of efficient tools for targeted genome modifications. Only during the past 5 years, a substantial number of different targeted genome-modified animals, including livestock and primates, have been produced for addressing issues in animal production or human medicine. It is reasonable to expect that our knowledge of functional genomics and applications of these tools will be greatly increased in the near future.

Among these designer nucleases, the CRISPR-Cas9 system is leading this advance and is currently the most convenient and cost-effective gene editing tool. Future investigation of the biochemistry and structure of the CRISPR-Cas system will help us design Cas9 mutants with enhanced performance, including improved specificity. Indeed, recent introduction of enhanced specificity of Cas9 is based on the elucidation of the crystal structure of the sgRNA-Cas9-target DNA complex^[88].

Acknowledgements The work was funded by the One Hundred Talents Program of Zhejiang University.

Compliance with ethics guidelines Shaohua Wang, Kun Zhang, and Yuning Dai declare that they have no conflict of interest or financial conflicts to disclose.

This article is a review and does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Niemann H, Kues W A. Application of transgenesis in livestock for agriculture and biomedicine. *Animal Reproduction Science*, 2003, 79(3–4): 291–317
2. Liu X, Wang Y S, Guo W J, Chang B H, Liu J, Guo Z K, Quan F S, Zhang Y. Zinc-finger nickase-mediated insertion of the lysostaphin gene into the beta-casein locus in cloned cows. *Nature Communications*, 2013, 4: 2565
3. Wall R J, Powell A M, Paape M J, Kerr D E, Bannerman D D, Pursell V G, Wells K D, Talbot N, Hawk H W. Genetically enhanced cows resist intramammary *Staphylococcus aureus* infection. *Nature*

- Biotechnology*, 2005, **23**(4): 445–451
4. Proudfoot C, Carlson D F, Huddart R, Long C R, Pryor J H, King T J, Lillico S G, Mileham A J, McLaren D G, Whitelaw C B, Fahnenkrug S C. Genome edited sheep and cattle. *Transgenic Research*, 2015, **24**(1): 147–153
 5. Luo J, Song Z, Yu S, Cui D, Wang B, Ding F, Li S, Dai Y, Li N. Efficient generation of myostatin (MSTN) biallelic mutations in cattle using zinc finger nucleases. *PLoS ONE*, 2014, **9**(4): e95225
 6. Yang P, Wang J, Gong G, Sun X, Zhang R, Du Z, Liu Y, Li R, Ding F, Tang B, Dai Y, Li N. Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin. *PLoS ONE*, 2008, **3**(10): e3453
 7. Yu S, Luo J, Song Z, Ding F, Dai Y, Li N. Highly efficient modification of beta-lactoglobulin (BLG) gene via zinc-finger nucleases in cattle. *Cell Research*, 2011, **21**(11): 1638–1640
 8. Brophy B, Smolenski G, Wheeler T, Wells D, L'Huillier P, Laible G. Cloned transgenic cattle produce milk with higher levels of β -casein and κ -casein. *Nature Biotechnology*, 2003, **21**(2): 157–162
 9. Prather R S, Lorson M, Ross J W, Whyte J J, Walters E. Genetically engineered pig models for human diseases. *Annual Review of Animal Biosciences*, 2013, **1**: 203–219
 10. Cooper D K, Ekser B, Ramsoondar J, Phelps C, Ayares D. The role of genetically engineered pigs in xenotransplantation research. *Journal of Pathology*, 2016: **238**(2): 288–299
 11. Bosze Z, Baranyi M, Whitelaw C B. Producing recombinant human milk proteins in the milk of livestock species. *Advances in Experimental Medicine and Biology*, 2008, **606**: 357–393
 12. Hammer R E, Pursel V G, Rexroad C E Jr, Wall R J, Bolt D J, Ebert K M, Palmiter R D, Brinster R L. Production of transgenic rabbits, sheep and pigs by microinjection. *Nature*, 1985, **315**(6021): 680–683
 13. Palmiter R D, Brinster R L, Hammer R E, Trumbauer M E, Rosenfeld M G, Bimberg N C, Evans R M. Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature*, 1982, **300**(5893): 611–615
 14. Lai L, Kolber-Simonds D, Park K W, Cheong H T, Greenstein J L, Im G S, Samuel M, Bonk A, Rieke A, Day B N, Murphy C N, Carter D B, Hawley R J, Prather R S. Production of α -1,3-galactosyl-transferase knockout pigs by nuclear transfer cloning. *Science*, 2002, **295**(5557): 1089–1092
 15. Cong L, Ran F A, Cox D, Lin S L, Barretto R, Habib N, Hsu P D, Wu X B, Jiang W Y, Marraffini L A, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 2013, **339**(6121): 819–823
 16. Mali P, Yang L H, Esvelt K M, Aach J, Guell M, DiCarlo J E, Norville J E, Church G M. RNA-guided human genome engineering via Cas9. *Science*, 2013, **339**(6121): 823–826
 17. Jackson D A, Symons R H, Berg P. Biochemical method for inserting new genetic information into DNA of Simian virus 40: circular Sv40 DNA molecules containing lambda phage genes and galactose operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 1972, **69**(10): 2904–2909
 18. Gordon J W, Scangos G A, Plotkin D J, Barbosa J A, Ruddle F H. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 1980, **77**(12): 7380–7384
 19. Jaenisch R. Germ line integration of moloney leukemia virus: effect of homozygosity at the m-mulV locus. *Cell*, 1977, **12**(3): 691–696
 20. Jaenisch R, Mintz B. Simian virus 40 DNA sequences in DNA of healthy adult mice derived from preimplantation blastocysts injected with viral DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 1974, **71**(4): 1250–1254
 21. Krimpenfort P, Rademakers A, Eyestone W, van der Schans A, van den Broek S, Kooiman P, Kootwijk E, Platenburg G, Pieper F, Strijker R, de Boer H. Generation of transgenic dairy cattle using 'in vitro' embryo production. *Biotechnology*, 1991, **9**(9): 844–847
 22. Niemann H. Transgenic pigs expressing plant genes. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, **101**(19): 7211–7212
 23. Hofmann A, Kessler B, Ewerling S, Weppert M, Vogg B, Ludwig H, Stojkovic M, Boelhaue M, Brem G, Wolf E, Pfeifer A. Efficient transgenesis in farm animals by lentiviral vectors. *EMBO Reports*, 2003, **4**(11): 1054–1058
 24. Kubisch H M, Larson M A, Eichen P A, Wilson J M, Roberts R M. Adenovirus-mediated gene transfer by perivitelline microinjection of mouse, rat, and cow embryos. *Biology of Reproduction*, 1997, **56**(1): 119–124
 25. Webster N L, Forni M, Bacci M L, Giovannoni R, Razzini R, Fantinati P, Zannoni A, Fusetti L, Dalpra L, Bianco M R, Papa M, Seren E, Sandrin M S, Mc Kenzie I F, Lavitrano M. Multi-transgenic pigs expressing three fluorescent proteins produced with high efficiency by sperm mediated gene transfer. *Molecular Reproduction and Development*, 2005, **72**(1): 68–76
 26. Maga E A, Geoffrey Sargent R, Zeng H, Pati S, Zarling D A, Oppenheim S M, Collette N M B, Moyer A L, Conrad-Brink J S, Rowe J D, BonDurant R H, Anderson G B, Murray J D. BonDurant R H, Anderson G B, Murray J D. Increased efficiency of transgenic livestock production. *Transgenic Research*, 2003, **12**(4): 485–496
 27. Scherer S, Davis R W. Replacement of chromosome segments with altered DNA sequences constructed in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 1979, **76**(10): 4951–4955
 28. Thomas K R, Folger K R, Capecchi M R. High frequency targeting of genes to specific sites in the mammalian genome. *Cell*, 1986, **44**(3): 419–428
 29. Smithies O, Gregg R G, Boggs S S, Koralewski M A, Kucherlapati R S. Insertion of DNA sequences into the human chromosomal β -globin locus by homologous recombination. *Nature*, 1985, **317**(6034): 230–234
 30. Evans M J, Kaufman M H. Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 1981, **292**(5819): 154–156
 31. Schwartzberg P L, Goff S P, Robertson E J. Germ-line transmission of a c-abl mutation produced by targeted gene disruption in ES cells. *Science*, 1989, **246**(4931): 799–803
 32. Hall B, Limaye A, Kulkarni A B. Overview: generation of gene knockout mice. *Current Protocols in Cell Biology*, 2009. doi: 10.1002/0471143030.cb1912s44 (Published online in Wiley Online Library)
 33. Wilmut I, Schnieke A E, McWhir J, Kind A J, Campbell K H. Viable offspring derived from fetal and adult mammalian cells. *Nature*, 1997, **385**(6619): 810–813

34. Dai Y, Vaught T D, Boone J, Chen S H, Phelps C J, Ball S, Monahan J A, Jobst P M, McCreath K J, Lamborn A E, Cowell-Lucero J L, Wells K D, Colman A, Polejaeva I A, Ayares D L. Targeted disruption of the $\alpha 1,3$ -galactosyltransferase gene in cloned pigs. *Nature Biotechnology*, 2002, **20**(3): 251–255
35. Sedivy J M, Sharp P A. Positive genetic selection for gene disruption in mammalian cells by homologous recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 1989, **86**(1): 227–231
36. Vasquez K M, Marburger K, Intody Z, Wilson J H. Manipulating the mammalian genome by homologous recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 2001, **98**(15): 8403–8410
37. Kuroiwa Y, Kasinathan P, Matsushita H, Sathiyaselan J, Sullivan E J, Kakitani M, Tomizuka K, Ishida I, Robl J M. Sequential targeting of the genes encoding immunoglobulin-mu and prion protein in cattle. *Nature Genetics*, 2004, **36**(7): 775–780
38. Mansour S L, Thomas K R, Capecchi M R. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature*, 1988, **336**(6197): 348–352
39. Klymiuk N, Mundhenk L, Kraehe K, Wuensch A, Plog S, Emrich D, Langenmayer M C, Stehr M, Holzinger A, Kroner C, Richter A, Kessler B, Kurome M, Eddicks M, Nagashima H, Heinritz K, Gruber A D, Wolf E. Sequential targeting of CFTR by BAC vectors generates a novel pig model of cystic fibrosis. *Journal of Molecular Medicine*, 2012, **90**(5): 597–608
40. Klymiuk N, Blutke A, Graf A, Krause S, Burkhardt K, Wuensch A, Krebs S, Kessler B, Zakhartchenko V, Kurome M, Kemter E, Nagashima H, Schoser B, Herbach N, Blum H, Wanke R, Aartsma-Rus A, Thirion C, Lochmuller H, Walter M C, Wolf E. Dystrophin-deficient pigs provide new insights into the hierarchy of physiological derangements of dystrophic muscle. *Human Molecular Genetics*, 2013, **22**(21): 4368–4382
41. Segal D J, Carroll D. Endonuclease-induced, targeted homologous extrachromosomal recombination in *Xenopus* oocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 1995, **92**(3): 806–810
42. Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 1994, **91**(13): 6064–6068
43. Chouluka A, Perrin A, Dujon B, Nicolas J F. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 1995, **15**(4): 1968–1973
44. Bibikova M, Carroll D, Segal D J, Trautman J K, Smith J, Kim Y G, Chandrasegaran S. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Molecular and Cellular Biology*, 2001, **21**(1): 289–297
45. Kim H, Kim J S. A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics*, 2014, **15**(5): 321–334
46. Pavletich N P, Pabo C O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*, 1991, **252** (5007): 809–817
47. Bibikova M, Golik M, Golik K G, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics*, 2002, **161**(3): 1169–1175
48. Morton J, Davis M W, Jorgensen E M, Carroll D. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103** (44): 16370–16375
49. Foley J E, Maeder M L, Pearlberg J, Joung J K, Peterson R T, Yeh J R. Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. *Nature Protocols*, 2009, **4**(12): 1855–1867
50. Cui X, Ji D, Fisher D A, Wu Y, Briner D M, Weinstein E J. Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nature Biotechnology*, 2011, **29**(1): 64–67
51. Yu S, Luo J, Song Z, Ding F, Dai Y, Li N. Highly efficient modification of beta-lactoglobulin (BLG) gene via zinc-finger nucleases in cattle. *Cell Research*, 2011, **21**(11): 1638–1640
52. Hauschild J, Petersen B, Santiago Y, Queisser A L, Carnwath J W, Lucas-Hahn A, Zhang L, Meng X D, Gregory P D, Schwinzer R, Cost G J, Niemann H. Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**(29): 12013–12017
53. Whyte J J, Zhao J G, Wells K D, Samuel M S, Whitworth K M, Walters E M, Laughlin M H, Prather R S. Gene targeting with zinc finger nucleases to produce cloned eGFP knockout pigs. *Molecular Reproduction and Development*, 2011, **78**(1): 2
54. Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annual Review of Phytopathology*, 2010, **48**: 419–436
55. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 2009, **326** (5959): 1509–1512
56. Moscou M J, Bogdanove A J. A simple cipher governs DNA recognition by TAL effectors. *Science*, 2009, **326**(5959): 1501
57. Christian M, Cermak T, Doyle E L, Schmidt C, Zhang F, Hummel A, Bogdanove A J, Voytas D F. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, 2010, **186**(2): 757–761
58. Carlson D F, Tan W, Lillico S G, Stverakova D, Proudfoot C, Christian M, Voytas D F, Long C R, Whitelaw C B, Fahrenkrug S C. Efficient TALEN-mediated gene knockout in livestock. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, **109**(43): 17382–17387
59. Butler J R, Ladowski J M, Martens G R, Tector M, Tector A J. Recent advances in genome editing and creation of genetically modified pigs. *International Journal of Surgery*, 2015, **23**: 217–222
60. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, 1987, **169**(12): 5429–5433
61. Mojica F J M, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*,

- 2000, **36**(1): 244–246
62. Jansen R, Embden J D, Gaastra W, Schouls L M. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 2002, **43**(6): 1565–1575
 63. Mojica F J, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, 2005, **60**(2): 174–182
 64. Pourcel C, Salvignol G, Vergnaud G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*, 2005, **151**(3): 653–663
 65. Bolotin A, Quinquis B, Sorokin A, Ehrlich S D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, 2005, **151**(8): 2551–2561
 66. Tang T H, Bachelier J P, Rozhdestvensky T, Bortolin M L, Huber H, Drungowski M, Elge T, Brosius J, Huttenhofer A. Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proceedings of the National Academy of Sciences of the United States of America*, 2002, **99**(11): 7536–7541
 67. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero D A, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 2007, **315**(5819): 1709–1712
 68. Brouns S J, Jore M M, Lundgren M, Westra E R, Slijkhuys R J, Snijders A P, Dickman M J, Makarova K S, Koonin E V, van der Oost J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, 2008, **321**(5891): 960–964
 69. Marraffini L A, Sontheimer E J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, 2008, **322**(5909): 1843–1845
 70. Garneau J E, Dupuis M E, Villion M, Romero D A, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadan A H, Moineau S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 2010, **468**(7320): 67–71
 71. Deltcheva E, Chylinski K, Sharma C M, Gonzales K, Chao Y, Pirzada Z A, Eckert M R, Vogel J, Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 2011, **471**(7340): 602–607
 72. Yang H, Wang H Y, Shivalila C S, Cheng A W, Shi L Y, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*, 2013, **154**(6): 1370–1379
 73. Li W, Teng F, Li T D, Zhou Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nature Biotechnology*, 2013, **31**(8): 684–686
 74. Hai T, Teng F, Guo R F, Li W, Zhou Q. One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. *Cell Research*, 2014, **24**(3): 372–375
 75. Niu Y Y, Shen B, Cui Y Q, Chen Y C, Wang J Y, Wang L, Kang Y, Zhao X Y, Si W, Li W, Xiang A P, Zhou J K, Guo X J, Bi Y, Si C Y, Hu B, Dong G Y, Wang H, Zhou Z M, Li T Q, Tan T, Pu X Q, Wang F, Ji S H, Zhou Q, Huang X X, Ji W Z, Sha J H. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell*, 2014, **156**(4): 836–843
 76. Gilbert L A, Larson M H, Morsut L, Liu Z, Brar G A, Torres S E, Stern-Ginossar N, Brandman O, Whitehead E H, Doudna J A, Lim W A, Weissman J S, Qi L S. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, 2013, **154**(2): 442–451
 77. Cheng A W, Wang H, Yang H, Shi L, Katz Y, Theunissen T W, Rangarajan S, Shivalila C S, Dadon D B, Jaenisch R. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Research*, 2013, **23**(10): 1163–1171
 78. Tanenbaum M E, Gilbert L A, Qi L S, Weissman J S, Vale R D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, 2014, **159**(3): 635–646
 79. Qi L S, Larson M H, Gilbert L A, Doudna J A, Weissman J S, Arkin A P, Lim W A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, 2013, **152**(5): 1173–1183
 80. Thakore P I, D'Ippolito A M, Song L Y, Safi A, Shivakumar N K, Kabadi A M, Reddy T E, Crawford G E, Gersbach C A. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nature Methods*, 2015, **12**(12): 1143–1149
 81. Chen B, Gilbert L A, Cimini B A, Schnitzbauer J, Zhang W, Li G W, Park J, Blackburn E H, Weissman J S, Qi L S, Huang B. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, 2013, **155**(7): 1479–1491
 82. Whitworth K M, Lee K, Benne J A, Beaton B P, Spate L D, Murphy S L, Samuel M S, Mao J D, O'Gorman C, Walters E M, Murphy C N, Driver J, Mileham A, McLaren D, Wells K D, Prather R S. Use of the CRISPR/Cas9 system to produce genetically engineered pigs from *in vitro*-derived oocytes and embryos. *Biology of Reproduction*, 2014, **91**(3): 78
 83. Petersen B, Niemann H. Molecular scissors and their application in genetically modified farm animals. *Transgenic Research*, 2015, **24**(3): 381–396
 84. Wang T, Wei J J, Sabatini D M, Lander E S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, 2014, **343**(6166): 80–84
 85. Kim S, Kim D, Cho S W, Kim J, Kim J S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, 2014, **24**(6): 1012–1019
 86. Ran F A, Hsu P D, Lin C Y, Gootenberg J S, Konermann S, Trevino A E, Scott D A, Inoue A, Matoba S, Zhang Y, Zhang F. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, 2013, **154**(6): 1380–1389
 87. Fu Y F, Sander J D, Reyon D, Cascio V M, Joung J K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology*, 2014, **32**(3): 279–284
 88. Slaymaker I M, Gao L, Zetsche B, Scott D A, Yan W X, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science*, 2016, **351**(6268): 84–88
 89. Zetsche B, Gootenberg J S, Abudayyeh O O, Slaymaker I M, Makarova K S, Essletzbichler P, Volz S E, Joung J, van der Oost J, Regev A, Koonin E V, Zhang F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 2015, **163**(3): 759–771