

Application of the CRISPR/Cas9 Gene Editing Technique to Research on Functional Genomes of Parasites

Running Title: CRISPR/Cas9 in parasites

Yubao Cui ^a, LiLi Yu ^b

^a Department of Clinical Laboratory, The Third People's Hospital of Yancheng, Affiliated Yancheng Hospital, School of Medicine, Southeast University, Yancheng 224001, Jiangsu Province, P. R. China

^b Department of Laboratory Medicine, Yancheng Health Vocational & Technical College, Yancheng 224006, Jiangsu Province, P. R. China.

***Corresponding author:** Dr Yubao Cui, Department of Clinical Laboratory, The Third People's Hospital, Affiliated Yancheng Hospital, School of Medicine, Southeast University, No. 299 at Jiefangnan Road, Yancheng 224000, Jiangsu Province, P. R. China. Email: ybcui1975@hotmail.com.

ABSTRACT

The clustered regularly-interspaced short palindromic repeats (CRISPR) structural family functions as an acquired immune system in prokaryotes. Gene editing techniques have co-opted CRISPR and the associated Cas nucleases to allow for the precise genetic modification of human cells, zebrafish, mice, and other eukaryotes. Indeed, this approach has been used to induce a variety of modifications including directed insertion/deletion (InDel) of bases, gene knock-in, introduction of mutations in both alleles of a target gene, and deletion of small DNA fragments. Thus, CRISPR technology offers a precise molecular tool for directed genome modification with a range of potential applications; further, its high mutation efficiency, simple process, and low cost provide additional advantages over prior editing techniques. This paper will provide an overview of the basic structure and function of the CRISPR gene editing system as well as current and potential applications to research on parasites.

Keywords: CRISPR/Cas9 systems; gene-editing technique; parasites; parasitic disease; diagnostics.

1. Introduction

Parasites pose a serious threat to both human and animal health. In fact, parasitic disease has restricted livestock development, and therefore economic advancement, in many parts of the world [1]. Advances in the sequencing and understanding of the genomes of parasites—including *Plasmodium* spp., *Schistosoma mansoni*, *Trypanosoma cruzi*, *T. brucei*, *Leishmania* spp., and *Theileria parva* [2, 3]—have allowed the study of functional genomes to play an increasingly important role in the science of parasitology. Genome sequence information will also enable the investigation of the relationship between a parasite and its host through both structural and functional genomics. Understanding these interactions could lead to new diagnostic methods for parasitic diseases and the development of vaccines and antiparasitic agents.

Genome editing techniques are important tools in modern biological research. One emerging genome editing technique uses clustered regularly interspaced short palindromic repeats (CRISPR) and the DNA endonuclease system of the associated Cas protein (CRISPR/Cas system). The CRISPR/Cas-based techniques provide new methods by which to manipulate the DNA of a variety of organisms and will enable detailed studies on the functional genomes of parasites. The CRISPR/Cas system is an acquired immune system found in bacteria and archaeobacteria [4-9]. This RNA-mediated immunity is used to fight invading viruses and plasmids by copying and cleaving the exogenous genetic material [6, 7]. Since the discovery of this system

in prokaryotes [4, 5], CRISPR and a specific endonuclease, Cas9, have been demonstrated to cut any DNA strand *in vitro* and have been used for efficient genome editing in a variety of eukaryotes, including *Drosophila*, yeast, mice, human cells, rats, zebrafish, bacteria, nematode, and rice [10-16]. This review focuses on the basic structure, mechanism, and application of the CRISPR/Cas9 system in parasites, and its prospects in future research.

2. The principle of the CRISPR/Cas system

CRISPR sequences were first discovered in *Escherichia coli* K12 cells [4]. Later work uncovered similar repeat structures in a wide range of bacteria and archaeobacteria [5, 8, 9]. It was initially unclear what role these conserved repeats might play in the prokaryotic genome [4, 5, 8, 9], but further studies discovered a function for CRISPR in acquired immunity [6, 7]. The system consists of tandem repeats of a common sequence with short spacers of variable sequence [6, 16]. These repeats are associated with nearby genes encoding nuclease or helicase proteins, called CRISPR-associated, or Cas proteins [6, 7, 16, 17]. The Cas proteins recognize, capture, and incorporate invading genetic material (and, in some cases, self-DNA) into spacers [6, 7, 18]. The CRISPR sequence is later transcribed into small RNAs that, together with Cas proteins, recognize foreign genetic material and mark it for degradation [6, 7], thereby eliminating the invading plasmid or virus.

The CRISPR system is divided into three types based on the nearby genes encoding

Cas proteins (for review, see [19]). The simplest system is Type II, in which only one effector Cas protein, rather than a complex of Cas proteins, is needed to perform the targeting and nuclease functions. Although a second Cas, Cpf1, was recently described as a putative Type II effector, Cas9 has been widely studied within this system as the major Type II effector protein [20]. Cas9 has two nuclease structural domains: an HNH nuclease structural domain within the protein and a RuvC-like structural domain located at the N-terminus. Each structural domain cuts both strands of the target DNA molecule [21]. Cas9 can process and produce CRISPR-RNA (crRNA) and, together with the trans-activating crRNA (tracrRNA), this complex can target and cleave exogenous DNA [19, 21].

Although much remains to be discovered regarding the underlying mechanisms of the CRISPR/Cas system, it can be roughly divided into three phases: adaptation, expression, and interference [19] (Figure 1). During the first phase, protospacers in the exogenous DNA are integrated between two repeats of the host genome at the CRISPR locus 5'-terminus [18-20]. In the second phase, CRISPR transcribes and produces precursor non-coding RNA (pre-crRNA). The long-chain pre-crRNA is then processed into short-chain crRNA. In the type II system, Cas9 plays an important role in this step [20]. The short-chain crRNA, together with tracrRNA, form a double-stranded secondary structure [19, 21] to guide Cas-mediated cleavage. Finally, during the third phase, binary complexes of crRNA and tracrRNA bind with the Cas protein. These complexes recognize specific loci and cut the double-stranded

exogenous DNA to produce double-stranded breaks; the Cas protein can be re-programmed to cut various targets based on the sequence of the short guide RNA (sgRNA), typically derived from the crRNA/tracrRNA [19, 21]. This system is highly amenable to engineering that allows it to specifically target nucleic acid sequences, making it a versatile tool for gene editing.

3. Application of the CRISPR/Cas system in *Plasmodium*

To date, the CRISPR/Cas technique has been applied in studies of *Plasmodium* spp. [22-24]. *Plasmodium* spp. are parasites transmitted via mosquitos and are the causal agent of malaria, which kills approximately 500,000 people annually [25]. Both gene editing and directional knockout (in non-essential genes) have been achieved using the CRISPR/Cas9 system [22-24], highlighting the versatility and potential application of this approach.

In one study, the *Streptococcus pyogenes* Cas9 endonuclease and sgRNA produced using T7 RNA polymerase were able to efficiently edit portions of the *P. falciparum* genome. Specifically, the authors targeted genes encoding knob-associated histidine-rich protein (kahrp) and erythrocyte-binding antigen 175 (EBA-175), resulting in transgenic parasites with high gene disruption frequencies [23]. In another study, Cas9 was used to produce DNA double-strand breaks in a specific locus in *P. yoelii*, with subsequent repair through homologous recombination. Genes with knockout efficiency up to 100% and cutting-in and allele replacement up to 22-45%

were achieved through providing homologous repair templates, generating targeted deletions, knocking-out, and replacing genes of parasites at multiple nucleotides [24]. Successful applications of the CRISPR/Cas technique to induce single-locus editing and multi-locus editing provide strong technical support for future studies on genome structure and gene family knockouts.

4. Application of CRISPR/Cas system in other parasites

Significant advancements have been made in genome editing of *Leishmania*, *Trypanosoma cruzi*, and *Toxoplasma gondii*, as well as vectors like the mosquito and water flea. Application of the CRISPR/Cas9 system in *Leishmania* has been successful using the Cas9 endonuclease to control the DHFR-TS promoter and the U6sn RNA promoter and terminator were used to produce the sgRNA [26]. Similarly, a more recent study generated loss-of-function insertion and deletion mutations in *L. donovani* [27]. These initial studies demonstrate proof-of-concept and open the door to future functional investigations.

In *Trypanosoma cruzi*, which causes Chagas disease, high-capacity repeat knockouts using CRISPR/Cas9 led to decreased expression of the β -galactofuranosyl glycosyltransferase family of enzymes, by targeting all 65 known members of the gene family with as few as three rounds of transfection [28]. Further, CRISPR/Cas9-mediated gene knockouts of *Pfr1*, *Pfr2*, and *Gp72* revealed roles for their gene products in flagellar attachment and cell motility [29].

In *Toxoplasma gondii*, the CRISPR/Cas9-based system has been used to guide destruction of RNA genes. In particular, gene editing was applied to knock out the *Rop18* gene and then use the mutant to transform type I GT1 cells [30]. This approach therefore enabled reverse genetics in a strain that was previously amenable only to forward genetic screens.

Beyond studying or altering the parasites themselves, the CRISPR/Cas9 system can enable gene editing in common vectors. Application of this technique in vectors is valuable as a means to uncover vector-parasite interactions, and to potentially introduce genetic changes that block the parasite's life cycle. The first advances in this have been achieved in vectors that transmit viruses rather than parasites. For example, in a transgenic mosquito which expressed enhanced cyan fluorescent protein (ECFP) and DsRed, the ECFP sequence was targeted for destruction using the CRISPR/Cas9 system. In four G1 pools, individuals still expressed DsRed, but no longer expressed ECFP. The study achieved a knockout efficiency of 5.5% [31]. Alternatively, Kistler et al. found that the CRISPR/Cas9 system can integrate exogenous gene sequences of mosquitos in a precise, targeted way, and that sgRNAs play an important role in investigating directional change of fatal genes of mosquitos [32]. Gantz et al. developed a Cas9-mediated synthetic system known as a gene drive, which passes a malaria-resistance gene onto both the male and female offspring of the mosquito *Anopheles stephensi* [33, 34]. Mosquitos transmit not only malaria but also

chikungunya, yellow fever, and dengue viruses; thus, the technology that has been demonstrated for virus-bearing species is likely transferable to those that transmit parasites. In addition, gene editing has been successful in water fleas, which transmit several parasites. When injecting Cas9 mRNA and eyeless-targeting RNA into the eggs of water fleas, 18-47% of the surviving larvae showed eye abnormalities as juveniles. After these juveniles matured, 8.2% produced offspring which showed deformities resulting from eyeless gene mutations [35].

The CRISPR technique achieves directional change of gene transcription in *Leishmania*, *Trypanosoma cruzi*, *Toxoplasma gondii*, water fleas, and mosquitos, and thus can readily be adapted for studies on parasite genome structure and function.

5. Conclusions

The CRISPR/Cas9 technique has been applied in a variety of biological studies since its development in early 2013. Compared to traditional transgenic techniques, the CRISPR/Cas9 system is more efficient and more accurate. It is easier to perform than the TALENs/ZFNs-based gene editing technique, and has higher knockout efficiency and more accurate gene editing. These advantages greatly reduce the off-target rate [19, 21, 36]. Further, the technique is relatively simple, low cost, and provides new possibilities for research related to genetic engineering. As such, the CRISPR/Cas9 technique will make a significant contribution to functional genomic research on parasites.

In early 2014, genome-scale knockout screening based on the CRISPR/Cas system was completed in human cells. The lentiviral transmission of genome-scale CRISPR/Cas9m knockout (GeCKO) of 18,080 targeting genes was suitable for positive and negative selective screening in human cells [37]. This finding holds significance for the application of this genome-scale screening platform to parasites. For example, to find the receptor of a specific protein in a certain parasite, sgRNA banks of each gene can be covered for genome-scale design to obtain individual organisms with corresponding genes knocked out. Subsequent protein screening could be performed to identify receptor-related genes.

The CRISPR/Cas9 system is in the initial stage of application to genome editing. It has primarily been used for studies of basic features of model organisms. However, to achieve wide application of this technique, further refinement is needed. Particular areas of optimization include resolving the off-target and unnecessary mutation problems, applying this technique successfully to the large and complex genomes of parasites, and improving the efficiency of homologous recombination and directional insertion of a gene segment.

With continued improvements to the CRISPR/Cas system, it could play an important role in genome-scale functional annotation of parasites. In the near future, this technique may be widely used in follow-up studies on the transcriptome and proteome

of parasites. Such studies could allow better understanding of genome function, expression, and regulatory mechanisms. Data from these studies could provide the experimental basis for applications related to clinical testing, early diagnosis, and disease prevention.

Conflicts of interest

The authors state that they have no Conflict of Interest (COI).

Acknowledgments

This work was supported by National Sciences Foundation of China (NSFC 30060166, NSFC81001330, NSFC31272369, NSFC31572319).

References

- [1] Food and Agriculture Organization of the United Nations. Animal Health Control.
<http://www.fao.org/ag/againfo/programmes/en/lead/toolbox/Tech/18Hyg.htm>.
Accessed on 14 October 2015.
- [2] A.O. Lau, An overview of the Babesia, Plasmodium and Theileria genomes: a comparative perspective, *Mol Biochem Parasitol*, 164(2009)1-8.
- [3] D.F. Wirth, Biological revelations, *Nature*, 419(2002)495-496.
- [4] Y. Ishino, H. Shinagawa, K. Makino, M. Amemura, A. Nakata, Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product, *J Bacteriol*, 169(1987)5429-5433.
- [5] F.J. Mojica, C. Diez-Villasenor, E. Soria, G. Juez, Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria, *Mol Microbiol*, 36(2000)244-246.

- [6] S.J. Brouns, M.M. Jore, M. Lundgren, E.R. Westra, R.J. Slijkhuis, A.P. Snijders, et al., Small CRISPR RNAs guide antiviral defense in prokaryotes, *Science*, 321(2008)960-964.
- [7] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, *Science*, 337(2012)816-821.
- [8] R. Jansen, J.D. van Embden, W. Gaastra, L.M. Schouls, Identification of a novel family of sequence repeats among prokaryotes, *OMICS*, 6(2002)23-33.
- [9] F.J. Mojica, C. Ferrer, G. Juez, F. Rodriguez-Valera, Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning, *Mol Microbiol*, 17(1995)85-93.
- [10] A.E. Friedland, Y.B. Tzur, K.M. Esvelt, M.P. Colaiacovo, G.M. Church, J.A. Calarco, Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system, *Nat Methods*, 10(2013)741-743.
- [11] W.Y. Hwang, Y. Fu, D. Reyon, M.L. Maeder, S.Q. Tsai, J.D. Sander, et al., Efficient genome editing in zebrafish using a CRISPR-Cas system, *Nat Biotechnol*, 31(2013)227-229.
- [12] J.E. DiCarlo, J.E. Norville, P. Mali, X. Rios, J. Aach, G.M. Church, Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems, *Nucleic Acids Res*, 41(2013)4336-4343.
- [13] D. Li, Z. Qiu, Y. Shao, Y. Chen, Y. Guan, M. Liu, et al., Heritable gene targeting in the mouse and rat using a CRISPR-Cas system, *Nat Biotechnol*, 31(2013)681-683.
- [14] S.W. Cho, S. Kim, J.M. Kim, J.S. Kim, Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease, *Nat Biotechnol*, 31(2013)230-232.
- [15] J. Miao, D. Guo, J. Zhang, Q. Huang, G. Qin, X. Zhang, et al., Targeted mutagenesis in rice using CRISPR-Cas system, *Cell Res*, 23(2013)1233-1236.

- [16] E. Hartenian, J.G. Doench, Genetic screens and functional genomics using CRISPR/Cas9 technology, *FEBS J*, 282(2015)1383-1393.
- [17] R. Jansen, J.D. Embden, W. Gaastra, L.M. Schouls, Identification of genes that are associated with DNA repeats in prokaryotes, *Mol Microbiol*, 43(2002)1565-1575.
- [18] A. Bolotin, B. Quinquis, A. Sorokin, S.D. Ehrlich, Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin, *Microbiology*, 151(2005)2551-2561.
- [19] J. van der Oost, E.R. Westra, R.N. Jackson, B. Wiedenheft, Unravelling the structural and mechanistic basis of CRISPR-Cas systems, *Nat Rev Microbiol*, 12(2014)479-492.
- [20] B. Zetsche, J.S. Gootenberg, O.O. Abudayyeh, I.M. Slaymaker, K.S. Makarova, P. Essletzbichler, et al., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system, *Cell*, 163(2015)759-771.
- [21] T. Karvelis, G. Gasiunas, V. Siksnys, Programmable DNA cleavage in vitro by Cas9, *Biochem Soc Trans*, 41(2013)1401-1406.
- [22] M.C. Lee, D.A. Fidock, CRISPR-mediated genome editing of *Plasmodium falciparum* malaria parasites, *Genome Med*, 6(2014)63.
- [23] J.C. Wagner, R.J. Platt, S.J. Goldfless, F. Zhang, J.C. Niles, Efficient CRISPR-Cas9-mediated genome editing in *Plasmodium falciparum*, *Nat Methods*, 11(2014)915-918.
- [24] C. Zhang, B. Xiao, Y. Jiang, Y. Zhao, Z. Li, H. Gao, et al., Efficient editing of malaria parasite genome using the CRISPR/Cas9 system, *MBio*, 5(2014)e01414-01414.
- [25] World Health Organization. Malaria Fact Sheet. Accessed from <http://www.who.int/mediacentre/factsheets/fs094/en/>. Accessed on 014 October 2015.
- [26] L. Sollelis, M. Ghorbal, C.R. MacPherson, R.M. Martins, N. Kuk, L. Crobu, et al., First efficient CRISPR-Cas9-mediated genome editing in *Leishmania* parasites, *Cell Microbiol*, 17(2015)1405-1412.

- [27] W.W. Zhang, G. Matlashewski, CRISPR-Cas9-Mediated Genome Editing in *Leishmania donovani*, *MBio*, 6(2015)e00861.
- [28] D. Peng, S.P. Kurup, P.Y. Yao, T.A. Minning, R.L. Tarleton, CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*, *MBio*, 6(2015)e02097-02014.
- [29] N. Lander, Z.H. Li, S. Niyogi, R. Docampo, CRISPR/Cas9-Induced Disruption of Paraflagellar Rod Protein 1 and 2 Genes in *Trypanosoma cruzi* Reveals Their Role in Flagellar Attachment, *MBio*, 6(2015)e01012.
- [30] B. Shen, K.M. Brown, T.D. Lee, L.D. Sibley, Efficient gene disruption in diverse strains of *Toxoplasma gondii* using CRISPR/CAS9, *MBio*, 5(2014)e01114-01114.
- [31] S. Dong, J. Lin, N.L. Held, R.J. Clem, A.L. Passarelli, A.W. Franz, Heritable CRISPR/Cas9-mediated genome editing in the yellow fever mosquito, *Aedes aegypti*, *PLoS One*, 10(2015)e0122353.
- [32] K.E. Kistler, L.B. Voss hall, B.J. Matthews, Genome engineering with CRISPR-Cas9 in the mosquito *Aedes aegypti*, *Cell Rep*, 11(2015)51-60.
- [33] V.M. Gantz, N. Jasinskiene, O. Tatarenkova, A. Fazekas, V.M. Macias, E. Bier, et al., Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*, *Proc Natl Acad Sci U S A*, 112(2015)E6736-6743.
- [34] S. Reardon, Welcome to the CRISPR zoo, *Nature*, 531(2016)160-163.
- [35] T. Nakanishi, Y. Kato, T. Matsuura, H. Watanabe, CRISPR/Cas-mediated targeted mutagenesis in *Daphnia magna*, *PLoS One*, 9(2014)e98363.
- [36] T.R. Sampson, S.D. Saroj, A.C. Llewellyn, Y.L. Tzeng, D.S. Weiss, A CRISPR/Cas system mediates bacterial innate immune evasion and virulence, *Nature*, 497(2013)254-257.
- [37] O. Shalem, N.E. Sanjana, E. Hartenian, X. Shi, D.A. Scott, T.S. Mikkelsen, et al., Genome-scale CRISPR-Cas9 knockout screening in human cells, *Science*, 343(2014)84-87.

Figure Legends

Figure 1. The three stages of CRISPR-derived immunity

The CRISPR/Cas (clustered regularly-interspaced short palindromic repeats/CRISPR-associated proteins) system comprises Cas proteins and the CRISPR locus, which consists of repetitive sequences (“repeats”, black diamond) that are separated by variable sequences (“spacers”, colored squares). In general, the CRISPR/Cas systems have been classified into three main types (type I, type II, type III) on the basis of the nearby genes encoding Cas proteins. Figure 1 is based on the mechanism of action of the type II CRISPR/Cas system, in which Cas9 plays an important role. CRISPR/Cas system involves three general phases: adaptation, expression, and interference. During adaptation, the exogenous DNA is fragmented and a protospacer is selected and integrated as a new spacer (purple square) at the leader-proximal end of the CRISPR locus. In the expression stage, these repeats and spacers are transcribed and then processed into mature CRISPR RNAs (crRNAs). Finally, during the interference stage, these mature crRNAs form CRISPR-Cas complexes with Cas proteins for targeting and cleaving the exogenous DNA. These complexes recognize invading DNA through the detection of a conserved protospacer adjacent motif (PAM), which is needed for spacer acquisition and interference.

Figure 1

