The CRISPR-Cas9 System: A Powerful Tool for Genome Engineering and Regulation

Yongmin Yan and Daoyan Wei*
Department of Gastroenterology, Hepatology & Nutrition, University of Texas M.D. Anderson Cancer, Houston, USA

Keywords: CRISPR; Cas9; Genome engineering

Abbreviations: CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats; Cas9: CRISPR associated (Cas) 9; sgRNA: Single-guide RNA

In this post-genome era of advanced high-throughput DNA/RNA sequencing technologies, information may no longer be a bottleneck to understand and tackle complicated genetic diseases such as cancer. What is still lacking, however, is an efficient, reliable, and easy tool to precisely modify the cellular genome for functional genome annotation, disease modeling, and possibly even corrective gene therapy.

Recently, an efficient, RNA-guided, site-specific DNA cleavage tool, CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats), and the CRISPR-associated (Cas)9 system has been developed from the Streptococcus pyogenes type II CRISPR adaptive immune system [1] and has attracted much attention for its potential to transform genome engineering and regulation.

In bacteria and archaea, CRISPR loci usually consist of three components: a cluster of cas genes and two non-coding RNA elements, trans-activating CRISPR RNA (tracrRNA) and a characteristic array consisting of repetitive sequences flanking unique spacer sequences (Figure 1A). Each spacer is derived from invading plasmid DNA. Transcription of the array yields individual CRISPR RNAs (crRNAs, consisting of spacer-repeat fragments), which localizes the crRNA: tracrRNA: Cas9 complex to target DNA where the effector Cas9 nuclease cuts both strands of DNA (double-strand breaks, DSBs) that matches the crRNA, and consequently, leads to the inactivation of invading DNA [2-4]. In mammalian and other cells, CRISPR-Cas induced DSBs can be repaired through two endogenous mechanisms: the non-homologous end joining (NHEJ) method which is generally used for the creation of a frameshift deleterious mutation, while the homology directed repair (HDR) is preferred for the introduction of a specific point mutation or addition of genes of interest. This precision targeting feature of the CRISPR-Cas9 system is of great interest for the study of biological processes [5].

What makes the CRISPR-Cas9 system even more attractive is the ease, high efficiency, and versatility of the technology. To simplify the CRISPR-Cas9 system, Jinek M et al. synthesized a single RNA chimera of dual-tracrRNA:crRNA (single-guide RNA, sgRNA), and successfully used it to direct sequence specific Cas9 double-strand DNA cleavage in a test tube in 2012 [6]. This study suggested the potential application of the CRISPR-Cas9 system for RNA-programmable genome engineering. In February 2013, two groups simultaneously demonstrated that the RNA-guided CRISPR-Cas9 system functions in both human and mouse cells and that multiplex editing of target genes is feasible upon introduction of multiple sgRNAs at the same time [6,7]. Shortly after these two milestone papers were published in Science, the CRISPR-Cas9 system was successfully used for genome modifications in other organisms such as plants [8-10], Caenorhabditis elegans [11-13], Drosophila [14,15], Zebrafish [10,15-17], and Xenopus tropicalis [18,19], suggesting that the CRISPR-Cas9 system may have broad applications in the biomedical sciences.

More recently, the CRISPR-Cas9 system was modified to create a more efficient, one-step, gene targeting technology. By co-injecting Cas9 mRNA and sgRNAs of interest into cells, Dr. Zhang’s group was able to simultaneously target five genes in mouse embryonic stem cells, and mice generated from zygotes co-injected with Cas9 mRNA and sgRNAs targeting Tet1 and Tet2 were shown to carry biallelic mutations in both genes with an efficiency of 80% [20]. This approach has much higher mutation efficiency and a much lower rate of off-target effects than the zinc-finger nuclease technique [21]. Similarly, reporter and conditional mutant mice were generated by this one-step co-injection of zygotes with Cas9 mRNA and different sgRNAs as well as DNA vectors [22]. Traditional generation of mice with multiple gene mutations requires careful breeding over many generations and may take 1-2 years. Therefore, this one-step approach to generating animals carrying mutations in multiple genes will greatly accelerate the in vivo study of gene functions and gene-gene interactions. Given that the CRISPR-Cas9 system’s sgRNAs are now much easier to make than proteins exploited in zinc finger and TALEN genome engineering technologies [23], it is possible to target virtually any gene using the CRISPR-Cas9 system, and a genome-wide resource of unique sgRNAs that target human exons is now available [7].

An additional modification of Cas9 has generated new applications of the CRISPR-Cas9 system. The CRISPR-Cas9 system’s genome editing function depends on the nuclease activity of Cas9, which cuts both strands of target DNA. Two nuclease domains have been identified in Cas9 protein, with the Cas9 HNH nuclease domain responsible for cleavage of the complementary strand, and the Cas9 RuvC-like domain responsible for cleavage of the non complementary strand (Figure 1B) [5]. Mutagenesis inactivation of its nuclease activity (so called dCas9, Cas9N, or Cas9nuclease-null) [24] retains Cas9’s RNA guided homing and DNA binding ability. The sgRNA-dCas9 complex’s specific DNA binding interferes with transcriptional elongation and RNA polymerase binding and transcription factor binding, and results in the silencing of gene expressions in bacteria, yeast, and human cells [25] [26]. This CRISPR interference (CRISPRi) is efficient and reversible and can simultaneously suppress multiple target genes [25]. Additionally, fusion of dCas9 with distinct regulatory domains resulted in specific, robust, or even inducible transcriptional activation or suppression of gene expressions in human and yeast cells [25,26,27,28], guided by sgRNAs, this system may help to map or perturb regulatory elements at.

*Corresponding author: Daoyan Wei, Department of Gastroenterology, Hepatology & Nutrition, University of Texas M.D. Anderson Cancer Center, Houston, TX77030, USA, Tel: 713-834-6683; E-mail: dwei@mtdanderson.org

Received October 27, 2013; Accepted November 01, 2013; Published November 06, 2013


Copyright: © 2013 Yan Y, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
the genome wide level. Similarly, engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) was established to isolate specific genome regions retaining molecular interactions. In enChIP, an antibody against a tag(s) fused to dCas9 is used for immunoprecipitation. [29]. Combined with mass spectrometry (MS) analysis, genomic loci-associated proteins have been identified using this enChIP-MS method [30]. These results demonstrate that the CRISPR-Cas9 system can be used as a general tool for precise gene regulation or to dissect the chromatin structure of genomic regions of interest.

Certainly, CRISPR-Cas9 technology is not limited to genome engineering and regulation. Functional genomic screening is a potential application of the CRISPR-Cas9 system when combined with sgRNA libraries. It is also anticipated that CRISPR-Cas9 technology may be used for therapeutic interventions to correct genetic disorders or modify endogenous protein expression levels [24].

To harness the technology of the CRISPR-Cas9 system, there are a number of issues that need to be considered, for example, how to select a target site and how to eliminate off-target effects. The ideal genome-engineering tool should be easily programmable to target the desired sequence within the genome without any off-target cleavage. The main advantage of CRISPR-Cas9 system is that it uses RNA as a guide for target recognition, which is easy to program and, theoretically, could be engineered to guide the crRNA:tracrRNA:Cas complex to any DNA target complementary to the crRNA. However, crRNA guided target selection is restricted by the protospacer adjacent motif (PAM) sequence, which is absolutely required for crRNA binding and cleavage by Cas9 nuclease, which has RuvC and HNH endonuclease domains that cut both strands of DNA, respectively.

Off-target effect is a critical issue in genome engineering, particularly for therapeutic application. In initial studies, significant off-target effects were observed using of the CRISPR-SpCas9 system, which may have been caused by the mismatch tolerance of SpCas9 and/or the unoptimized dosages of SpCas9 and sgRNA [6,32-34]. Longer PAM sequences should increase specificity and decrease the number of cutting sites within an interest region of genome, thus decreasing the off-target effect. The use of certain Cas9 orthologs that recognize different bacterial strains. For example, the PAM sequence required for SpCas9 (Streptococcus pyogenes) is 5’-NGG, whereas for NmCas9 (Neisseria meningitides) it is 5’-NNNNGATT or 5’-NNNNNGCTT [31].

References