CRISPR-CAS9: An Overview of Modern Genetic Editing

Mary Shannon Byers*

Department of Biological Sciences, Tennessee State University, Nashville, USA

*Corresponding Authors: Mary Shannon Byers, Department of Biological Sciences, Tennessee State University, 3500 John A. Merritt Blvd., Nashville, TN 37209, USA, E-mail: maryshannonbyers@yahoo.com

Received: 01 June 2019; Accepted: 11 June 2019; Published: 01 July 2019

Abstract
Gene editing is important for the development of highly productive, disease-resistant animals in agriculture, for the generation of scientific model organisms, and for biomedical research purposes. CRISPR-Cas9 editing processes involve DNA sequence modifications to targeted sites of the genome and can be used to silence undesirable traits in different organisms. Although there were concerns over off-target mutations involved in some of its earlier applications, the Type II CRISPR-Cas9 system has been demonstrated to produce mutations only at the targeted genomic site without inducing mutations at other functional regions of the genome. The CRISPR-Cas9 system has been utilized in various food production applications and in a variety of biomedical research studies on animal subjects and is considered safe for use in plants, non-human animals and primates, as well as for the therapeutic study of human diseases.

Keywords: Genetic; CRISPR; TALEN; DNA; Enzyme

Abbreviations: CRISPR-Clustered regularly interspaced palindromic repeats; gRNA-Guide RNA; sgRNA-Single guide RNA; PAM-Protospacer adjacent motif; crRNA-CRISPR RNA; tracrRNA-Trans-activating crRNA; U6-U6 RNA polymerase III promoter; siRNA-Short (small) interfering RNA; ZFN-Zinc finger nuclease; TALEN-Transcription activator-like effector nuclease

1. Introduction
In agriculture, gene editing is important for the development of highly productive disease-resistant animals, as well as for the generation of scientific model organisms for research purposes [1]. Editing processes involve DNA sequence modifications, such as insertions and deletions, to targeted sites of the genome. Some common editing technologies include the zinc finger nuclease (ZFN) [2, 3] and transcription activator-like effector nuclease (TALEN) systems [4-6], and now the modern Type II CRISPR-Cas9 system [7, 8]. ZFNs are restriction enzymes
created by the fusion of a zinc finger DNA-binding domain to a DNA-cleavage domain. TALENs are similarly generated by fusing a transcription activator-like effector DNA-binding domain to a DNA-cleavage domain. TALENs contain a central repeating domain, approximately 34 amino acid residues in length, with approximately 1.5 to 33.5 amino acid repeats [9].

The CRISPR-Cas9 system is a little more complex and offers more efficient editing compared to ZFN and TALEN systems. CRISPR-Cas technology is based on the adaptive immune defense mechanism of prokaryotic organisms. CRISPR is an acronym for clustered regularly interspaced palindromic repeats. Cas is short for CRISPR-associated genes and enzymes. The Cas9 enzyme contains two important domains, the n-terminus nuclease domain, and a second domain located within the center of the Cas9 enzyme. When the sgRNA binds to Cas9, Cas9 undergoes a conformational shape change, which then positions its domains at an angle to induce double-stranded breaks at the target site, three to four nucleotides upstream of the PAM sequence.

The ZFN, TALEN and CRISPR-Cas9 systems each induce targeted, double-stranded breaks to DNA, resulting in the DNA repair processes of homologous recombination and non-homologous end joining and can each be used for high frequency targeted gene replacement and mutagenesis. However, of these gene editing systems, CRISPR-Cas9 technology is the most highly effective and simplest to use and is believed to produce the least toxic effect on the organism being edited [10, 11]. For this reason, CRISPR-Cas9 technology is becoming increasingly more popular for use in agriculture and in biomedical research.

2. CRISPR-Cas adaptive immunity

The CRISPR-Cas adaptive defense system acts as a protective genetic memory bank to shield prokaryotic organisms against invasive DNA elements, which could become lethal to their systems. CRISPR-Cas immunity occurs via the uptake and integration of exogenous, invading DNA fragments to the organism’s CRISPR loci [12, 13]. The DNA fragment is transcribed to a short interfering gRNA segment, which forms a complex within the Cas enzyme by complementary base pairing with a Cas tracrRNA scaffold. In prokaryotic systems, the tracrRNA scaffold is a short, trans-encoded RNA segment important for the maturation of crRNA and for providing immune support against lysogenic phages [14]. This complex enables the Cas enzyme to recognize and destroy specific sequences from invasive bacteriophages, plasmids and other exogenous DNA elements [15, 16]. The CRISPR locus exists in roughly 45% of bacterial and 84% of archaeal organisms [17].

Prokaryotic CRISPR-Cas adaptive immunity consists of three stages: adaptation, expression and interference [15, 18, 19]. During the adaptation phase, fragmented DNA protospacers from viruses and plasmids are acquired into the repetitive CRISPR array of the prokaryotic immune system. The newly-acquired protospacers provide a unique sequence “memory bank,” enabling the organism to recognize subsequent invasions. Most CRISPR-Cas editing systems recognize a specific PAM motif to target DNA [20, 21], dependent upon the prokaryotic organism of the system. The expression phase is considered the biogenesis of crRNA. During the expression phase, the CRISPR
array is subject to transcription processing of pre-crRNA, the precursor transcript, to produce mature, small interfering crRNA. During the interference phase, the mature crRNA directs Cas enzymes to target and cleave specific homologous invading nucleic acids [22, 23]. In nature, these defense mechanisms allow prokaryotic organisms to target and silence invading DNA elements to protect their systems against future attacks. The three stages of immunity are believed to be associated with the many distinct Cas enzymes containing varying, individual functions.

3. CRISPR-Cas core

The CRISPR region is associated with a diverse array of Cas genes and enzymes [24]. The more common CRISPR-associated enzyme, Cas9, is an RNA-guided DNA endonuclease enzyme which cleaves near the center portion of a phosphodiester bond of the target DNA site, the restriction site. The DNA strand is unevenly cleaved, creating complementary, single-stranded sticky ends, which may be paired and re-joined by DNA ligase to connect to the phosphodiester bonds of the DNA fragments. There are roughly 93 known Cas protein families associated with the many CRISPR-Cas systems of the various prokaryotic organisms [24]. The Cas enzymes are categorized based on their sequence similarities. Eleven of the 35 families make up the Cas core, which consists of Cas protein families 1-11. The CRISPR-Cas systems are divided into Class I, which utilizes a multiple complex of discrete Cas proteins to degrade foreign nucleic acids, and Class II, which incorporates one large Cas protein for nucleic acid degradation. Each Class consists of three types with 19 subtypes. Class 1 contains Type I, Type III and Type IV, and Class 2 consists of Type II, Type V and Type VI. Each type, and most subtypes, are regulated by a unique signature Cas gene. Cas1 and Cas2 are conserved proteins of the adaptation phase of the prokaryotic immune system and are universal across each type and subtype. Cas3, Cas9 and Cas10 are universal across each type.

Cas1 and Cas2 are involved in the adaptation process of CRISPR-Cas systems throughout the majority of prokaryotic organisms [25]. Cas1 is a metal-dependent integrase enzyme featuring an α-helical fold [26]. Cas1 has the ability to cleave linear single-stranded DNA, single-stranded RNA and short double-stranded DNA fragments, as well as branched DNA substrates [27], assisting the Cas1 enzyme to mediate the insertion of new protospacers into the CRISPR array by targeted cleavage at specific sites [28, 29]. Cas1 forms a complex with Cas2 to generate unevenly cleaved double-stranded fragments of the CRISPR array near the leader repeat. Cleavage activity is followed by the rejoining of the 5'-end of the repeating strand to the 3'-end of the newly-incorporated protospacer sequence [25]. The Cas1-Cas2 complex undergoes a conformational change upon binding of the protospacer, creating a flat surface ranging from a Cas1 dimer across a Cas2 dimer and then making contact with a second Cas1 dimer. The short 23 base pair fragment is flanked by tyrosine-22 from the first and second Cas1 dimers, allowing for protospacer recognition [30, 31]. Stable integration of the protospacer requires a supercoiling of target DNA, and that the inserted protospacer sequence contains 3'-OH ends to allow for a nucleophilic attack on the minus strand of the first repeating sequence [29]. Cas1 may form interactions with exonuclease enzymes recB, recC and RuvB, suggesting that Cas1 may play a dual role in DNA repair mechanisms [27]. The role of Cas2 is less understood in comparison to Cas1. Cas2 is a homologue of the mRNA interferase toxins [32-35], and the active mutant site of the
Cas2 enzyme has been shown to acquire new protospacers [31]. However, although Cas2 is a major portion of the Cas1-Cas2 complex, Cas2 may not be required for enzymatic activity and does not influence intrinsic sequence specificity [28, 35].

Cas3 is associated with bacterial organisms such as *Streptococcus thermophilus* and *E. coli* and archaeal organisms like *Methanocaldococcus jannaschii*. Cas3 is a member of Class I, Type I. Cas3 acts as single-stranded DNA and single-stranded RNA endonuclease and exonuclease enzyme with ATPase activity, coupled to the unwinding of DNA-DNA and RNA-DNA duplexes. This activity occurs at the free single-stranded 3’ ends and is greater for DNA compared to RNA. Cas3 ATPase-helicase domain may serve as a motor protein, assisting in the CRISPR-associated activity of targeting foreign DNA. The Cas4 protein is a single-stranded DNA exonuclease enzyme and comes from archaeal organisms like *Sulfolobus solfataricus* and *Pyrobaculum calidifontis*. Cas4 is a member of Class II, Type IIB. Catalytic exonuclease cleavage occurs in the 5’ to 3’ direction and yields 5’-OH and nucleoside 3’-phosphate groups. Cas4 has magnesium-dependent endonuclease activity on circular single-stranded DNA. It can also unwind double-stranded DNA without ATP. Cas4 may contain a RecB nuclease domain and may play a role in capturing new viral DNA sequences for incorporation into the host genome.

Cas5, from bacterial organisms such as *Gardnerella vaginalis* and *Myxococcus xanthus*, belongs to Class I, Type IA. Cas6 is from archaeal organisms like *Halofex volcanii*, *Pyrococcus furiosus* and *Halofex mediterranei*, as well as bacterial organisms such as *Mycobacterium tuberculosis*. Cas7 has been found in bacteria such as Myxococcus xanthus and in archaeal organisms such as *Halofex volcanii*, *Thermoproteus tenax* and *Sulfolobus solfataricus*. These Cas enzymes may assist in the processing and stabilization of pre-CRISPR RNA to individual, mature CRISPR RNA units. The Cas8 proteins belong to Class I. Cas8 can be found in bacterial systems such as *Corynebacterium ulcerans*. Cas8a belongs to Type IA. Cas8a2 is found in archaeal organisms like *Thermoproteus tenax*. Cas8b belongs to Type IB and is found in bacteria such as *Campylobacter curvus*. Cas8c is among Type IC and has been found in *Streptococcus pyogenes*. Each Cas8 enzyme contains a subunit interference module important for recognizing and targeting invading DNA with a protospacer adjacent motif, or PAM, sequence.

The Cas10 isoforms are members of Class I. Cas10d belongs to Type ID and has a homologous domain to the nucleic acid polymerase and nucleotide cyclase palm domain. Another Cas10 enzyme belongs to Type III and is a homolog of Cas10d. Cas10c belongs to Type IIIIC. Its function has not yet been determined. Cas9 is a member of the Class II, Type II system. During the expression stage of immunity, pre-crRNA binds to a single, multidomain Cas9 enzyme, or to a multi crRNA-effector complex. The pre-crRNA is transcribed to mature crRNA by an endonuclease subunit of the complex [36]. In addition to Cas9 activity, another mechanism of RNA processing involves bacterial ribonuclease III (RNase III) and transactivating crRNA (tracrRNA) [14]. During the interference stage, mature crRNA remains bound to Cas9, giving it the ability to recognize and cleave foreign DNA and RNA. Endonucleases HNH and RuvC produce double-stranded breaks in DNA or may create single-stranded breaks individually.
4. TYPE II CRISPR-Cas9 System

The more popular Type II CRISPR-Cas9 system is based on the Cas9 enzymes of prokaryotic organisms such as *S. pyogenes* and *S. aureus*. The Type II system is beneficial for scientific research due to the Cas9 enzyme’s ability to cleave nearly any sequence complementary to a sgRNA strand. The Type II system has been successfully applied across a variety of disciplines from genomic regulation to epigenome engineering, and ranging in areas from agriculture to gene therapy applications [37]. Genomic editing with CRISPR-Cas9 technology can be accomplished with the utilization of a viral vector system to target PGCs or somatic cell lines and introduce exogenous genes into the organism, which could produce gene-edited transgenic animals [38].

5. CRISPR-Cas in Science

Although there were concerns over off-target mutations involved in some of its earlier applications [39-41], the Type II CRISPR-Cas9 system has been demonstrated to produce mutations only at targeted genomic sites without inducing mutations at other functional regions of the genome [42]. The CRISPR-Cas9 system has been utilized in various food production applications and in a variety of biomedical research studies on animal subjects and is considered safe for use in plants [43-45], non-human animals [46] and primates, as well as for the therapeutic study of human diseases. Some of its most significant contributions to biomedical research include the use of CRISPR-Cas9 to correct genetic mutations in skeletal muscle cells of mdx mouse models expressing Duchenne muscular dystrophy [47], as well as a study in rhesus monkeys, which could ultimately lead to translational research in humans [42]; the excision of the proviral genome HIV-1 from infected cells, which could provide a cure for HIV-infected patients [48]; genome editing of myosin binding protein C3 (MYBPC3) mutations for the treatment of hypertrophic cardiomyopathy [49]; and knockout of integrin a5 (ITGA5) from triple negative breast cancer cells for the treatment of metastatic triple negative breast cancer [50]. With the assistance of a modified CRISPR-Cas9 vector system, a research team at Vanderbilt University Medical Center demonstrated that particular single nucleotide polymorphisms (SNPs) associated with the expression of mouse *Arntl2*, a gene that encodes a circadian rhythm transcription factor, had a protective effect on the production of metastatic lesions in mouse models. In human breast cancer cases, those SNPs were associated with disease-free survival [51].

6. CRISPR-Cas9 Toxicity

Recent concerns involve CRISPR-Cas9’s associated risk of toxicity to human cells and the damaging effects of on-target mutations. Two recent studies have demonstrated that gene editing with CRISPR-Cas9 may trigger a cancer response in transfected cells. A team of researchers from the Karolinska Institute in Sweden showed that CRISPR-Cas9 editing of immortalized human retinal pigment epithelial cells generated a p53-mediated response, resulting in cell cycle arrest and selection against cells with functional p53 pathway. They determined that p53 inhibition prevented the cell cycle’s response to damage, leading to increased homologous recombination with a donor template [52]. Another team with the Novartis Institutes for Biomedical Research in Cambridge, MA had similar findings. The Novartis team revealed that a high transfection efficiency rate of CRISPR-Cas9 edited human pluripotent stem cells resulted in toxic double-stranded Cas9 breaks, which killed most of the human pluripotent
stem cells. The researchers concluded that genome editing induces a p53 and TP53 response, resulting in the toxicity [53]. Both groups advised that transfected cells should be monitored for p53 function [52, 53]. Another recent study revealed that CRISPR-Cas9 editing induced extensive damage from on-target copy number variations, including large deletions, insertions and inversions [54].

In addition to concerns over off-target mutations, another concern is the potentially harmful mutations that could be introduced through the disease-repair process, which could mediate repair in all cells, even in healthy ones [55], as well as the risk of human embryonic germline modifications on future generations [56]. There are additional concerns that CRISPR-Cas9 could potentially be exploited for non-therapeutic uses, such as for the editing of human embryos for the purpose of creating “designer babies” [56, 57].

7. Conclusions
Additional considerations involve changes made to an organism's metabolism, growth rate and response to environmental factors, which could have detrimental impacts on the surrounding environment. According to the USDA website, Congress has set up regulations to ensure the safety of genetically modified products based on the characteristics of the product and its intended use. The Food and Drug Administration, Animal and Plant Health Inspection Service and the Environmental Protection Agency oversee that all genetically modified products are safe for consumer use, safe for the environment and safe for use by farmers [58]. While the safety of human genome editing has not been yet been fully established [55], if ultimately proven to be safe, CRISPR-Cas9 protocols could be used in future studies to treat various genetic conditions [59].

References


This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license 4.0.