CRISPR-Cas: Removing Boundaries of the Nature

Nicat Cebrailoglu1, Ali Burak Yildiz1, Ozlem Akkaya1, Yelda Ozden Ciftci1

1Gebze Technical University, Department of Molecular Biology and Genetics, Kocaeli, Turkey

ORCID IDs of the authors: N.C. 0000-0002-2770-5049; A.B.Y. 0000-0002-9982-7109; O.A. 0000-0003-0473-1417; Y.O.C. 0000-0002-9799-3648


ABSTRACT

The CRISPR-Cas 9 system, which is known as a natural way of bacteria to defend against phage infection and plasmid transfer, has been re-purposed as a RNA-guided DNA targeting strategy for genome editing. Together with the advances gained in DNA sequencing technology, this platform opened a new era in molecular biology since its recognition was specified by 20-nt single-guide RNA which made technique easier, efficient and simple for application in any organism. Thus, many studies have discussed and performed the applications of CRISPR-Cas systems on different organisms for genome editing. Moreover, targeted gene regulations, epigenetic modulation, chromatin imaging and manipulation could also be applied with this system. Besides all its potential promising aspects, this tool might have some side effects like off-target mutations. In addition, unexpected results have also been reported after some gene editing applications. Thus, this review provides a brief history of gene editing tools together with the overview of the latest applications, regulations and ethical/structural aspects of the CRISPR Cas system.

Keywords: Gene editing, Gene therapy, Meganucleases, TALEN, ZFN, New breeding technologies

INTRODUCTION

Investigations on natural protection ways of bacteria against phages resulted in tremendous turning points in recombinant DNA technology. Starting with the discovery of restriction enzymes in the late 1970s that enabled scientists to manipulate DNA in test tubes (1), it allowed many opportunities of genetic manipulations in many organisms including bacteria, plants, animals and even humans. The key developments on the precise alteration of DNA in living eukaryotic cells, which is termed as “gene editing” (GE), started with Rothstein’s report in 1983 on yeast cells. Afterwards, Smithies and co-workers (1985) followed by Capecchi (2) demonstrated that it was possible to incorporate an exogenous copy of DNA into the mammalian cells genome through homologous recombination (HR). Although these studies resulted in the characterization of functional roles of many genes in model organisms, they have the following limitations, such as i) the rate of spontaneous integration was too low (1 in 10^3-10^6 cells, Capecchi, 1989), ii) type and the state of the cell affected the integration rate, iii) the possibility of random integration of exogenous copy to undesired site was similar or even higher than the target site (3).

To overcome these obstacles, scientists started to use different approaches among which the construction of a double-strand break (DSB) at a target site provided the best alternative for the elevation of targeted gene integration frequency. Thus, natural rare cutting meganucleases (i.e., I-SceI) and then re-engineered ones were utilized to achieve targeted DSBs. Even though these attempts resulted in some improvements, these enzymes had several disadvantages listed in Table 1. Afterwards, zing fingers, that are zinc ion-regulated small proteins that recognize and bind a 3 bp DNA sequence (4), were fused with the DNA cleavage domain of the Fok I endonuclease which is isolated from Flavobacterium okeanokoites to create a programmable
nucleases (5). Zinc finger nucleases (ZFN) then increased the capability to edit genomes at the targeted sites enabling the usage of this technique for therapeutic applications (6). Likewise, Fok I DNA cleavage domain is also combined with TALE modules in order to be utilized as an effective programmable nuclease (TALEN, transcription activator-like effector nucleases) (7). In contrast to ZFN that recognize a 3 bp, TALE proteins from Xanthomonas bacteria can recognize one single base. By using these nucleases, a DSB can be introduced in any site of the genome with known recognition sites of the DNA-binding domains. However, it should also be noted that TALEN nuclease sites require T before the 5’-end of the target sequence which could limit its application.

Although “Nature Methods” announced ZFN and TALEN as the method of the year for precise GE tools (https://www.nature.com/articles/nmeth.1852.pdf), their disadvantages indicated in Table 1 made researchers seek alternative approaches such as the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated protein 9 (Cas9) system. CRISPR-Cas9 system, which is known as a natural way of bacteria to defend against phage infection and plasmid transfer, has been re-purposed as a RNA-guided DNA targeting strategy for genome editing and opened a new era in molecular biology since its recognition was specified by 20-nt single-guide RNA (sgRNA) (8).

This recent platform not only mimics the natural trans-activating CRISPR RNA (tracrRNA) and CRISPR RNA (crRNA) structure, but also in contrast to ZFN and TALEN, there is no need for tedious protein engineering of DNA-recognition domains for each target site which make the design easy, simple to use and efficient (9).

### CRISPR-Cas: STRUCTURE AND FUNCTION

CRISPR-Cas systems are based on two molecules, guide RNA (gRNA) and Cas protein which are responsible of binding on a specific target at the genome and cutting the target point, respectively. The most applied and discussed one is the CRISPR-Cas9 system, which causes a DSB at the specific target area on the genome and silencing of the gene. This DSB might be repaired by either homology directed repair (HDR) or non-homologous end joining (NHEJ) systems. During the repairing of the gene via these two different types of systems, some indels might occur. Except triple insertion or deletions, other nucleotide changes cause frameshift mutations on the gene. Even though Streptococcus pyogenes (SpCas9) is the most used and studied one, there are more types of Cas proteins which are parts of immune metabolisms of archaea and bacteria. In order to function properly, those proteins require some specific short nucleotide sequences near the target area, which is called the protospacer adjacent motif (PAM) that may vary between different Cas proteins (10).

<table>
<thead>
<tr>
<th>Table 1: The advantages and disadvantages of gene editing tools.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene Editing Techniques</strong></td>
</tr>
<tr>
<td><strong>Meganucleases</strong></td>
</tr>
<tr>
<td><strong>ZFNs</strong></td>
</tr>
<tr>
<td><strong>TALENs</strong></td>
</tr>
<tr>
<td><strong>CRISPSR/CAS9</strong></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th><strong>CRI</strong></th>
<th><strong>RISPO/CAS9</strong></th>
<th><strong>t</strong></th>
<th><strong>RIPVE</strong></th>
<th><strong>CASMER</strong></th>
<th><strong>OR</strong></th>
<th><strong>S</strong></th>
<th><strong>N</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>i</strong></td>
<td><strong>p</strong></td>
<td><strong>s</strong></td>
<td><strong>p</strong></td>
<td><strong>r</strong></td>
<td><strong>a</strong></td>
<td><strong>s</strong></td>
<td><strong>9</strong></td>
</tr>
</tbody>
</table>

---
Newly discovered CasX protein has a great potential of human GE since it has a small size and the transformation of this protein to human cells is much easier than other Cas proteins. In addition to that, unlike some Cas9 proteins, CasX proteins are not found in bacteria which live in the human body and CasX proteins do not have any common ancestors with Cas9 proteins. Because of this reason, human immune systems are not capable of showing a strong response against CasX proteins (11). Similar to CasX, Cas14 also has a small size and does not require PAM sequence to generate single strand DNA break (12). Another Cas protein is Cas12a, also known as Cpf1, causes sticky DSB which makes easier the gene insertion into target loci (13).

Depending on the purpose, the Cas protein might be used single or as a fusion protein. For instance, Cas9 protein with an inactivated catalytic side (dCas9) is fused with cytidine deaminase to convert cytidines to uridines without causing any DSB on the target side (14). This method might be used for opposite change, uridines to cytidines with adenosine deaminase (15). CRISPR-Cas system is also utilized for generating epigenetic changes. There are three types of epigenetic changes generated by the CRISPR-Cas system: i) single or multiple gRNAs can manipulate dCas9 fused to a VP64 transcriptional activation domain to enhance the expression of endogenous human genes (16), ii) dCas9-KRAB fusions block the binding or progressing of DNA polymerase, resulting in repression of transcription (17), iii) targeted DNA methylation editing by using dCas9-TET1 catalytic domain fusions (18).

Since DNA editing causes permanent changes on the genome, off-target brings some huge risks for the future of the organism. This situation gives researchers a reason to use the CRISPR-Cas system on RNA editing to prevent these risks, as these types of changes stay at the transcriptional level. For RNA editing, researchers use two types of Cas proteins, Cas13a and Cas13b. Cas13a is used for mRNA degradation, on the other hand dCas13b-ADAR fusion is used for base editing on mRNA (19). There are several methods for delivering components of the CRISPR-Cas system to an organism. These methods are classified as physical (electroporation) (20), viral (adenovirus, lentivirus, tobacco rattle virus) (21-23) and non-viral methods (lipid nano particles, agrobacterium) (24, 25). After delivery, for the screening of the changes on cell or organisms, several methods are used. Screening of large-scale mutation requires sequencing or determination of the DNA band size with electrophoresis (26). For screening of small-scale mutation T7 endonuclease assay or restriction enzyme assay might be used (27, 28).

As discussed before, CRISPR is a mechanism of bacterial immune system against phages. On the other hand, phages also developed inhibitor proteins against this bacterial mechanism. Those proteins are called anti-CRISPR proteins, which allow us the control application of CRISPR and make clinical trials safer (29).

RECENT APPLICATIONS

In order to use this technology as a gene therapy tool, CRISPR needs to be delivered to the right cells in the human body. There are two different applications: in vivo and in vitro. Genome editing via cas-9 in vivo has been used to correct alleles associated with genetic diseases in animal models. In order to lead to cataract-free progeny, CRYGC gene, causing dominant-negative cataract, mutated by injecting Cas9 mRNA into the zygote of the mouse heterozygote and a sgRNA targeting only in the mutant allele (30). Furthermore, in case of this technique was applied to a mdx mouse which had a mutation in the gene encoding dystrophin, phenotypic correction was observed between 2% to 9100 percent when Cas9, sgRNA and donor template were injected into mouse zygotes (31). All of these studies showed promising advances in the treatment of genetic diseases.

Streptococcus pyogenes Cas9 (SpCas9) promises great potential for curing hereditary disorders including muscle dystrophy, HIV, vision disorders and many others. However, for all these applications to be possible, the dose and timing of SpCas9 activity should be adjusted to reduce the effects of off-targets. If SpCas9 activity can be controlled in these aspects; editing of DNA in model organisms can be successfully achieved. As an example, gene drives in genetically altered mosquitoes can prevent the spread of malaria and similar diseases transmitted by mosquitoes. The demand for the control of SpCas9 activity has raised a requirement for anti-CRISPR molecules. Although the anti-CRISPR proteins target SpCas9 are large and cannot pass through the cells, which can break down by proteases and cause the formation of an adverse immune reactions in the body. Whereas small molecule inhibitors are proteolytically stable and generally do not produce an immune response since they can diffuse through the cells. Future studies are needed to identify the mechanisms of action of the inhibitors on SpCas9: gRNA binding domains (32).

Although many studies have focused on the Cas9 protein so far; in recent years, the CRISPR-Cpf1 protein, also known as Cas12a, has been shown to be more effective than Cas9. As a matter of fact, companies like Mammoth Biosciences have already started using Cas12a technology. Patents containing the Cas12a-RNA complex are supported by the Berkeley and the Broad Institute. Recently CasX, which was originally discovered in Jennifer Doudna’s laboratory in 2017, is much smaller than other Cas proteins and has the ability to shade both Cas9 and Cas12a. Since it sources are from bacteria that are not found in humans, the human immune system is more likely to accept it than Cas9 (32, 33). Differently, the newly developed Cas13-based SHERLOCK, which targets RNA, allows us to diagnose multiple diseases with one test and gives us a hundred times more sensitive results. Following the bonding of Cas13 with the viral genome, Cas13 starts to cut free specific RNAs, and these RNA cuttings trigger the formation of signals.
Currently, CRISPR gene-editing technology has been started to be used for human clinical trials: β-thalassemia [Vertex Pharmaceutical/CRISPR Therapeutics], Cancer (melanoma, sarcoma, myeloma) [U Penn/Parker Institute] and HIV [Affiliated Hospital to Academy of Military Medical Sciences]. The earlier studies led by Feng Zhang (MIT) and George Church (Harvard University) showed that the CRISPR system could be used to edit eukaryotic mammalian cells, including human cells. Later, a lot of researches were performed in this field. In November 2018 the Chinese researcher He Jiankui made the world’s first genetically edited babies. He used CRISPR to mutate the gene called CCR5. Disabling this gene would prevent the HIV virus from entering and destroying Helper T cells. If everything had been gone as planned, children with an immune response to AIDS would have been born. However, it was shown that the growing CRISPR babies may face earlier deaths (average of 1.9 years) (34) a the genetic mutation that protects against HIV causes the babies to have a shorter life span.

Moreover, CRISPRi in which dCas9 is fused to a transcriptional repressor domain [Kruppel associated box (KRAB)] for repression of transcription and CRISPRa in which dCas9 is fused to a sequence (SunTag) containing multiple copies of the activator recruitment domain of the general control protein (GCN4) to activate transcription are also used to elucidate the non-coding genome (35-37). Genome editing by this method allows for efficient disorder of regulatory elements without causing DNA mutations. To sum up, dCas9-based methods enable to clarify the roles of regulatory sequences within the natural genomic structure and can elucidate long non-facilitating RNAs that can be altered by indels generated with Cas9 nucleases (38).

The chromatin structure modulates the genome. However, elucidation of the basis of this modification depends on a limited number of methods used to study chromatin-protein interactions. To identify proteins that interact with a specific genome locus; the chromatin may be precipitated with an antibody against a dCas9-tag fusion protein expressed together with gRNA targeting the desired DNA sequence, which is called ‘engineered DNA-binding molecule-mediated chromatin immunoprecipitation’ (enChIP). Later on, locus-associated proteins can be identified by mass spectrometry (39). enChIP is used in living cells for biochemical analysis of transcription and epigenetic regulation (40).

CRISPR-based tools can create a new guide RNA, enabling easier genome-wide screening. Many researchers have reported that screening with this method is more specific and more efficient than RNAi, and yields more robust and trustable results (41). In these experiments, sgRNA libraries and Cas9 cells are introduced into the cell, and selection of the treated cells was conducted according to those showing the targeted phenotypic result (42). Such screenings have been used to identify genes that are involved in cancer progression (43), drug resistance (44), immune response (45) susceptibility to bacterial toxins (46) and the emergence of other biomedically important phenotypes.

The ability to screen multiple loci in the human genome at the same time by performing a single experiment via Cas9 (47) enables the identification of complex cell signaling pathways, gene functions, drug targets for therapeutic purposes, and predicting drug side effects.

**PLANT CRISPR EDITING**

In plant biotechnology, CRISPR is used for both improving and gaining features on plants, such as yield and quality (48-52), herbicide tolerance (53, 54), biotic and abiotic tolerance (26, 55-57). Along with those, CRISPR is also used for functional genomic (58, 59) studies. High mutation frequency of CRISPR is especially important for creating homozygote mutant lines on polyploid plants such as wheat, potato and strawberry. In regards to potato, researchers established a single base change in the ALS gene of the tetraploid plants by using Cas9-Cytidine deaminase fusion, and made the plant resistant to the chlorsulfuron herbicide. This mutation prevents the acetolactate synthase enzyme from inhibition by chlorsulfuron binding (60). Furthermore, CRISPR has made it easy to target multiple genes in a single organism. For instance, researchers mutated different genes in cultivated tomatoes with a single CRISPR application to ensure recovery of stress tolerance since tomato has lost its tolerance to stress due to domestication for a long time. As a result, tomatoes with a bigger fruit size, number and nutritional value were obtained (61). In another study, the Fad 2.1 gene, responsible for the conversion of oleic acid to less stable linoleic acid, was knocked-out by Calyxt Inc. to increase the oleic acid content of the soybean and ultimately the shelf life of the soybean oil. This first commercial GE plant, which was developed for the first time with TALEN, was later achieved by the CRISPR technique (62).

**BACTERIAL GENOME EDITING WITH CRISPR**

Since there are other effective methods for genome editing in microorganisms such as HR; few studies have been reported on the development of CRISPR-Cas genome editing in various bacteria (i.e., E. coli, Cyanobacteria, Streptomyces, Riemerella anatipestifer, Clostridium, Corynebacterium, Bacillus, Salmonella, Pseudomonas putida, Lactobacillus casei) (63-72). Identification of strains (73), detection of natural or engineered immunity against mobile genetic elements (74, 75), manipulation of microbial consortium (76), and programmable transcriptional regulation (77) are some issues that have been tried to be solved using CRISPR. Moreover, patent studies in this field focused on the growth of microorganisms, preventing antibiotic resistance, biofuel production and enhanced synthesis of desired metabolites (78). The development of this method will enable efficient screening and selection of targeted mutations in microorganisms.

**ETHICS AND SAFETY REGULATIONS**

Since the first publication in 2012 (79) that reported CRISPR-Cas9 usage for genome editing; this method has been described by different names such as “revolutionary”, a “groundbreaking” and
“game changer”, since it provided the opportunity of crossing species boundaries. Naturally, this facility seemed to be very promising at first glance for many researchers.

Clinical trials using CRISPR system for efficient genome editing of various mammalian cells have already started and give promise to the treatment of some major diseases (80-83). In fact, these clinical applications also highlighted the presence of certain risks. Ihry et al. (84) revealed that DSBs generated by Cas9 could be toxic, and it created an obstacle for high genome-editing efficiency of CRISPR/Cas9 in human pluripotent stem cells. This study implied that using CRISPR in human cell lines increased the risk of cancer. Moreover, unexpected mutations resulting from CRISPR editing are another issue that needs improvement (85). Since these mutations can cause various genetic disorders or cancer, some social and ethical doubts about this genome engineering tool have appeared. Editing the unborn child to have the desired eyes or hair color (86), building an army with genetically edited soldiers (87) could be a few of the future applications of this unlimited technology. Despite all these possible risks and ethical considerations, the US and China are the countries that have allowed researchers to apply CRISPR editing on human CAR-T cells (88, 89). In addition to that, as it stated above, an illegal experiment was reported in China to make HIV resistant babies (90). However, it was later showed that HIV-resistant babies with CCR5 mutations were also sensitive to dangerous flu and West Nile Virus (91). All those present studies suggest that even though CRISPR is a very powerful technique, it is not always the first option in curing diseases.

Likewise, all these ethical and social aspects should also be discussed for gene-editing applications in agriculture. Since CRISPR provides gene editing without any DNA integration, in many countries, engineered plants using this technique have been accepted as a non-transgenic product that is allowed to enter the market freely without the need for regulation (92-97). Only the European Union decided that edited crops should be considered as GMOs (Table 2). However, in some European countries, field trials of altered plants using this approach are still ongoing. It is also expected that after “Brexit” the UK might remove the regulations for CRISPR-edited crops. Recently, 14 European Union (EU) countries have already made a call for updating the laws of GMOs according to New Plant Breeding Technologies (98,99).

CONCLUSIONS

Although CRISPR-Cas system offers tremendous opportunities for clinical and biotechnological applications, it might cause some unexpected results which reveals that the technique needs to be improved and further tested. Besides, ethical issues and regulatory aspects should also be discussed in scientific consortia to have a common decision on GE applications prior to large-scale clinical and field applications.
Peer-review: Externally peer-reviewed.


Conflict of Interest: The authors declare that they have no conflicts of interest.

Financial Disclosure: There are no funders to report for this submission

REFERENCES


Fujita T, Fujii H. Identification of proteins associated with an IFNγ-responsive promoter by a retroviral expression system for enChIP using CRISPR. PLOS ONE 9 2014; e103084.


86. Beaumont KA, Shekar SN, Cook AL, Sturm RA. Red hair is the null phenotype of MC1R. Hum Mutat 2008; 29(8): E88-E94.


