Crisprs/Cas9 May Provide New Method for Drug Discovery and Development

Zhabiz Golkar*, Luigenia Rochelle2 and Omar Bagasra3
1Department of Biology, School of Health and Natural Science, Voorhees College, Denmark
2Division of Academic and Student Affairs, Voorhees College, Denmark
3Division of Biology, South Carolina Center for Biotechnology at Claflin University, USA

*Corresponding author: Zhabiz Golkar, Assistant Professor, Department of Biology, School of Health and Natural Science, Voorhees College, Denmark. Tel: (803)-780-1500; E-mail: zgolkar@voorhees.edu

Rec date: Feb 05, 2016; Acc date: Mar 26, 2016; Pub date: Mar 28, 2016

Copyright: © 2016 Golkar Z, 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.

Abstract

The CRISPR/Cas9 (Clustered Regions of Interspersed Palindromic Repeats-Cas9), consists of a DNA-nuclease and a piece of RNA that homes in on a DNA sequence, enabling investigators to create precisely targeted mutations, corrections to mutations, or other gene modulation. CRISPR/Cas9 is an ancient anti-viral immune system found in bacteria and archaea. The immediate assumption would be that it is a primitive innate immune system like a restriction enzyme defense system known since 1970 and is the backbone of the gene cloning methods. Surprisingly, it is a sophisticated adaptive immune system very different from the somatic gene recombination, which is well-known and is found in higher vertebrate animals as T and B lymphocytes. This amazing newly discovered CRISPR has emerged as a magnum opus of programmable nuclease technology for the precise editing of the genome in cells. This new genome editing tool is much more robust to customize and optimize because the site selection for DNA cleavage is guided by a short sequence of RNA. Even though this tool still has some imperfections and suffers from some off-target effects, the CRISPR/Cas9 system has been widely and successfully applied as a biotechnology in a number of areas. This technology is being considered to edit defective genes in human embryos and to create specific DNA fragment insertions for correcting numerous genetic diseases. The following section is a brief history and development of the CRISPR system and shows its potential future applications. We believe that the readers will benefit greatly from the information of this newly discovered prokaryotic adaptive immunity system, and we believe that in the very near future this technology will be widely used in clinics and research.

Keywords: CRISPR; Anti-CRISPR; Cas; Cas9; crRNA; Drug discovery; Antibiotic resistance

Introduction

One hundred and fifty years ago, Gregor Mendel discovered the basic principles of heredity through experiments in his garden. Mendel’s observations became the foundation of modern genetics and the study of heredity. Within approximately 15 years, scientists achieved such impressive feats as sequencing the entire human genome and gaining understanding of the causes of many genetic diseases. However, despite the significant understanding of genetics and technology, the challenge of altering genomes precisely at the single gene level still persists. The development of efficient and reliable ways to make precise, targeted changes to the genome of living cells is a long-standing goal for biomedical researchers.

Genetic engineering has been applied to a diverse range of technologies, including drug development, gene therapy, molecular evolution and synthetic biology. Despite the predicted utility of a successful gene editing technique, many current methods such as zinc fingers nucleases (ZFN) and TALENs have confounding of low efficiency, time-consuming procedures, and lack of specificity for both model organisms and humans. In recent years, a new gene editing technique, derived from Streptococcus pyogenes, has generated considerable excitement and is a frontrunner for effective and successful gene editing. Clustered regularly interspaced short palindromic repeats (CRISPRs) are DNA loci that contain multiple, short, direct repetitions of base sequences which are inherited as a result of previous infections. After expression, RNA-guided gene-editing platform makes a small strand of RNA that interacts with invasive DNA. Cas genes code for proteins involved in CRISPR function makes use of a bacterial protein 9 and a systematic guide RNA to introduce a double stranded break at a specific location within the genome (Figure 1). In nature, the CRISPR/Cas system makes up the prokaryotic adaptive immune system by defending against infection by plasmids and phages through silencing of exogenous DNA invading material. Researchers can exploit this mechanism by introducing plasmids with Cas genes and specifically designed CRISPRs for an organism. This unique system fosters another genetic modification of the genome of most of the prokaryotic or eukaryotic organisms at a precise site. This article discusses the mechanism, applications, and future potential of the CRISPR/Cas System.

Discovery of CRISPR/Cas9

CRISPR were first described in 1987 as the bacterium Escherichia coli by Ishino lab [1], but at that time their function was not known. Ishino et al, discovered a group of 29-nucleotide repeats divided by non-repetitive short sequences in E. coli, which are now known as spacers. In 2000, similar repeats were determined in other bacteria and archaea, and were termed short regularly spaced repeats (SRSR) [2]. SRSR were renamed CRISPR in 2002 [3]. The authors also found a conserved set of genes generally associated with CRISPR repeats, and...
was named the Cas, or CRISPR-associated, genes. This family consists of genes encoding proteins such as DNA helicases (Cas3) and exonucleases (Cas4), which implies that Cas genes likely have an important role in DNA metabolism (Figure 2).

Figure 1: The arrangement of three major components of the Clustered regularly-interspaced short palindromic repeats (CRISPR) in Prokaryotic DNA.

In 2005, Bolotin [4] and two other independent studies showed that some CRISPR spacers are originated from phage DNA and extra chromosomal DNA such as plasmids [5,6]. Bolotin was studying the bacteria _S. thermophilus_, which had just been sequenced, revealing an unusual CRISPR locus [4]. Although the CRISPR array was comparable to previously reported systems, it lacked some of the known Cas genes and instead contained unique Cas genes, including one encoding a large protein they predicted to have nuclease activity, which is now known as Cas9. Furthermore, they noted that the spacers, which are homolytic to viral genes, all share a common sequence at one end. This sequence, the proto-spacer adjacent motif (PAM), is required for target recognition [6]. In 2006, Marakova et al. discussed the idea that CRISPR-based immunity could function through a mechanism analogous to RNAi [7]. In 2007, Barrangou and Horvath (food industry scientists at Danisco) and Moineau's group at Université Laval showed that spacer DNA could be used to alter the resistance of _S. thermophilus_ to phage attack [8]. _S. thermophilus_ is widely used in the dairy industry to make yogurt and cheese; these findings were important for scientists to answer a common problem in industrial yogurt making [8]. The experimental study of Horvath et al. showed that CRISPR systems are a prokaryotic adaptive immune system: they integrate new phage DNA into the CRISPR array, which allows them to fight off the next wave of attacking phage with similar genetic homologies [8].

Moreover, they study showed that sequence identity between the spacer and its match in the phage genome (the proto-spacer) is required for CRISPR immunity, that new repeat-spacer units are acquired upon phage challenge, and that Cas genes are necessary for CRISPR function. More recently, Marrafiﬁni et al. showed that CRISPR interference can limit plasmid conjugation in _Staphylococcus epidermidis_ demonstrating a broader role for CRISPR in the prevention of HGT in bacteria [9].

**CRISPR Mechanism**

One of the prokaryotic and archaea’s immune system that is based on a region of DNA called CRISPRs, is largely a stand-alone system that is capable of functioning in an individual cell, a necessity for organisms that often display unicellular behavior. The CRISPR-Cas system, which targets DNA or RNA as a way of protecting against viruses and other mobile genetic elements, can be found on both chromosomal and plasmid DNA, and the spacers are often derived from genetic sequences of viruses and plasmids. It is assumed that these viral genetic sequences were incorporated into the genomes of the host bacteria previously and now these viral genetic elements can be utilized to degrade the same or similar viral genes, providing the host cells with adaptive immunity. Interestingly, any time bacteria or archaea get infected with a new virus, the system adds new spacers so new viruses can be recognized. The spacers are used as recognition elements to fight matching virus genomes and destroy them.

CRISPR defense requires the presence of a set of CRISPR-associated or _Cas_ genes, generally found adjacent to the CRISPR, that code for proteins essential to the immune response. Since the genome is modiﬁed in the process of spacer acquisition, offspring inherit the protection. New spacers are usually added at one side of the CRISPR, making the CRISPR a chronological record of the viruses the cell and its ancestors have encountered previously.

The CRISPR-Cas mediated defense mechanism operates in three sequential phases. The first phase is to acquire the new invading viral sequence, known as “adaptation” or “acquisition”, which leads to insertion of new spacers in the CRISPR locus (Figure 2). In the second phase, the acquired sequence expresses the _Cas_ genes and transcribes the CRISPR into a long precursor CRISPR RNA (pre-cRNA). The pre-cRNA is subsequently processed into mature crRNA by Cas proteins and accessory factors. In the third and last phase, the target nucleic acid is recognized and destroyed by the combined action of crRNA and Cas proteins.

The length and the acquired sequence of repeats and the length of the acquired target sequences are well conserved within a CRISPR locus, but may vary between CRISPR in the same or different genomes. Repeat sequences are in the range of 21 bp to 48 bp, and spacers are between 26 bp and 72 bp. This is not surprising, however, since a specific transposon that is invading a cell may have similar sequences with multiple related and unrelated viruses, and adaptation of a broader sequence may protect the host against multiple invaders. The Cas proteins that are engaged in various defense functions in the whole CRISPR-Cas systems are a highly diverse group. Many have been identified to interact with nucleic acids; _e.g._ as nucleases, helicases and RNA-binding proteins. Therefore, the Cas1 and Cas2 proteins are
involved in adaptation and are virtually universal for CRISPR-Cas systems. Other Cas proteins are only associated with certain types of CRISPR-Cas systems.

It is believed that similar to the molecular immunity, the CRISPR-Cas systems probably originated from mobile genetic elements that frequently transfer horizontally, which also contributes to their high prevalence. The functions of CRISPR and CRISPR-associated (Cas) genes are critical in adaptive immunity in selected bacteria and archaea, empowering the organisms to respond to and eradicate invading genetic material. Three types of CRISPR mechanisms have been identified, of which type II is studied the most.

The four steps of CRISPR/Cas complex

Step 1: Spacers conforming to fragments of DNA from a phage/foreign DNA are integrated into CRISPR sequence. It is assumed that both Cas1 and Cas2 proteins participate in this process. At this step, sets of new spacers are added at the leading end of the CRISPR sequence.

Step 2: A noncoding sequence, rich in AT (Adenine and Thymine) at the leading end of CRISPR, acts as a promoter and as a result CRISPR loci are transcribed into pre-crRNA.

Step 3: Pre-crRNA have to be processed into crRNA. In this phase also, each piece of crRNA comprises of a single spacer between two half-repeats.

Step 4: Finally, crRNA is used to destroy foreign genetic material [10] (Figures 2 and 3).

The CRISPR/Cas system is similar in principle to the RNA interference (RNAi) used by eukaryotic cells; both use short RNA sequences to guide the destruction of foreign nucleic acids by enzymes. [11] However, these two systems employ entirely different sets of proteins. No homology has been found between CRISPR/Cas and RNAi [8,12]. Although the alternating spacer repeat structure of CRISPR sequences is highly conserved across all species, the protein complex system which uses crRNA to eradicate foreign genetic material is diverse [13]. In some cases, an array of highly combined complex systems of many different proteins are used to destroy foreign DNA [14]. (The statement as originally written was difficult to decipher; however, the revision is red is my guess of what may be the intended meaning). In some other cases, a single protein with a guide RNA may be sufficient to cleave DNA e.g. CRISPR-associated complex for antiviral defense CASCADE, CMR complex, and Cas9 [15].

The Different Types of CRISPR-Cas System

Three types of CRISPR mechanisms have been classified, of which type II the most is studied. Type I loci contain the Cas3 gene, which encodes a large protein with distinct helicase and DNase activities [16]. In addition to genes encoding proteins that probably form Cascade-like complexes with different compositions, [16] these complexes contain numerous proteins that have been included in the RAMP.
superfamily [17]. RAMP superfamily is linked to CRISPR loci which contains six genes, cmr, cmr2, cmr3, cmr4, cmr5 and cmr6 are present in a range of bacteria and archaea and not present an autonomous functional unit but always associated with one of CRISPR subtypes. In the Cascade complex, a RAMP protein with RNA endonuclease activity has been identified as the main enzyme that catalyzes the processing of the long spacer-repeat-containing transcript into a mature crRNA [16,17]. In most cases, the catalytic RAMP proteins (Cas6, Cas6e and Cas6f) do not belong to the most prevalent Cas5 or Cas7 families of RAMPs and are often encoded in the periphery of the respective operon (Figure 3A).

The type II CRISPR mechanism is unique compared to other CRISPR systems since it includes the ‘HNH'-type system and Cas9 is required for gene silencing [18]. During the destruction of target DNA, the HNH and Ruv C-like nuclease domains cut both DNA strands, thus generating double-stranded breaks (DSBs) at sites defined by a 20-nucleotide target sequence within an associated crRNA transcript. The HNH domain cleaves the complementary strand, while the RuvC domain cleaves the non-complementary strand [15,16,18]. Type II seems to be sufficient for generating crRNA and destroying the target DNA, in addition to the ubiquitous Cas1 and Cas2. Cas9’s function in both of these steps and relies on the presence of two nuclelease domains: a Ruv C-like nuclelease domain located at the amino terminus and an HNH-like nuclelease domain that resides in the mid-region of the protein [18,19]. Type II systems cleave the pre-crRNA through a mechanism that involves duplex formation between a tran-crRNA and part of the repeat in the pre-crRNA; the first cleavage in the pre-crRNA processing pathway subsequently occurs in this repeat region. This cleavage is catalyzed by the housekeeping, double-stranded RNA-specific RNase III in the presence of Cas9 [19] (Figure 3B).

The type III CRISPR-Cas systems contain polymerase and RAMP modules in which at least some of the RAMPs seem to be involved in the processing of the spacer-repeat transcripts, analogous to the Cascade complex. Type III systems can be further divided into subtypes III-A (Mtube or CASS6) and III-B (polymerase-RAMP module) [20,21] (Figure 3C).

**Figure 3:** Three different Types of CRISPR-Cas system; the CRISPR/Cas systems have been classified into three distinct types: A) type I, B) type II and C) type III. While types I and III are found in both bacteria and archaea, type II is unique to only bacteria. The bacterial type II CRISPR/Cas system is the most studied and best characterized of the three types in which Cas9 protein is the critical component. In this system, CRISPR loci are transcribed as a precursor CRISPR RNA (pre-crRNA) containing the full set of CRISPR repeats and embedded invader-derived sequences 47. A trans-activating crRNA (tracrRNA) binds to the repeat sequences of the pre-crRNA to form a duplex RNA, which is then cleaved by a double-stranded RNA-specific ribonuclease RNase III, the Cas9 protein.

**Anti-CRISPR Mechanisms**

Just as cells have developed numerous strategies to counter viruses, the viruses have developed counter measures to these strategies [22]. The most fundamental way for viruses to escape CRISPR-Cas activity is by random mutagenesis that affects key bases in the interface with the crRNA or the PAM recognition [23]. It has been determined in *P. aeruginosa* phages that encode several proteins affecting the activity of Type I-E and I-F systems. Functions of these proteins and expression of Cas proteins or crRNA are still unknown and probably interfere with activity of CRISPR-Cas system [23]. In general, the few phages that manage to infect do so by having picked up a new spacer targeting the host locus, and therefore demonstrating that the viruses can use the full adaptive potential of the CRISPR-Cas system [24].
CRISPR and evolution of antibiotic resistance pathogens

Horizontal gene transfer (HGT) is a major source of genetic variability for bacterial evolution [18]. Interestingly, it has been shown that the CRISPR/cas system may slow the spread of antibiotic resistance genes [25]. It has been proven that CRISPR systems have the ability to limit phage infection and plasmid conjugation [23,25]. It is not known yet if CRISPR is establishing an effective system against DNA transformation, although it was shown that CRISPRs can inhibit the electroporation of plasmid DNA [26]. Consequently, CRISPR systems inhibit at least two major routes of HGT and thus could have an important role in bacterial evolution.

In addition to phage DNA, the study showed [27] that the CRISPR/Cas system can destroy plasmids. Horizontal transfer of plasmids containing resistance genes is a common mechanism for the spread of antibiotic resistance [27]. If the CRISPR/Cas system targets plasmids containing resistance genes, transformation will not take place. Consequently, it may be feasible to manipulate the CRISPR/Cas system to slow the spread of antibiotic resistance genes in a clinical setting.

Furthermore, HGT is the major mechanism for the procurement of antimicrobial resistance genes or virulence gene factors in bacterial pathogens. The study of Golkar, et al. [28] on the threat of antibiotic resistant crisis revealed that the emergence of methicillin-resistant Staphylococcus aureus (MRSA) and vancomycin-resistant S. aureus (VRSA) are health-care issues in clinical settings [28]. MARSA and VRSA strains transfer antibiotic resistance genes by plasmid conjugation, which enhance fitness and pathogenicity of those bacterial strains [28]. S. aureus and S. epidermidis strains are the most common causes of nosocomial infections, and mobile genetic elements can spread from one species to the other.

CRISPR interference has been found to limit conjugation of the pG0400 plasmid from S. aureus to S. epidermidis in the laboratory and possibly constitutes a natural barrier to the spread of antimicrobial resistance [26,28]. Upon infection of the bacterial host, phages can undergo either lytic or lysogenic replication cycles. In the lysogenic cycle, a temperate phage integrates its genome into the bacterial chromosome, becoming an inheritable prophage. It has been long known that prophage-encoded genes have an important role in the virulence of pathogenic strains [29]. Conversely, strains that lack CRISPRs are poly-lysogens. Moreover, many CRISPR spacers match sequences of prophages that are integrated into other strains; that is, there is a mutually exclusive relationship between CRISPR spacers and prophages, which suggests that CRISPR immunity can prevent not only phage lysis but also lysogenesis. Therefore, CRISPR immunity against lysogenic bacteriophages may interfere with the spread of virulence factors among pathogens.

Finally, many of the virulence plasmids that are required for establishing a successful infection by a number of bacterial pathogens are believed to have diverged from conjugative plasmids [28,29]. Also, pathogenicity islands are flanked by transposable elements and therefore can transfer between different species by ‘hitch-hiking’ on conjugative plasmids and temperate phages [29]. The prevention of conjugation and phage infection by CRISPRs suggests a capacity for these loci to reduce the acquisition of genetic traits that allow bacteria to become virulent.

Gene knock-out application of CRISPR in eukaryotic organisms

Recent work has proven that the CRISPR/Cas9 system can be utilized for gene editing in a plethora of systems including yeast, rice, zebra fish, mice, and even humans [30]. Based on the type II CRISPR/ Cas9 mechanism, researchers have engineered an RNA chimera of tracrRNA-crRNA called single guide RNA (sgRNA), which is able to cause sequence-specific binding to double stranded DNA [31]. Cas9 nuclease is capable of interacting with this engineered sgRNA in order to specifically cleave dsDNA in regions complementary to the introduced sgRNA (Figure 4). The presence of double stranded breaks (DSB) in the DNA leads to activation of the DSB repair machinery for either non-homologous end joining (NHEJ) or the homology directed repair (HDR) pathway, which requires the presence of a repair template. The HDR mechanism copies the sequence of the template into the cut target sequence to repair the DSB. This method has been found to work at a high efficiency for genome editing in most eukaryotic model systems [28,31]. CRISPR could be used to delete harmful disease causing mutations in the human genome, especially in IVF embryos [29,31]. Also, this system could be used to knock out specific genes in model organisms in order to study genetic diseases, for example, by using the CRISPR/Cas9 mediated homologous recombination in intestinal stem cells for human stem cell therapy and HIV therapy. However, more work must be done to determine accuracy and to reduce the risks of the technique [32].

Development of CRISPR/Cas9 delivery systems

In order to exploit gene editing in vivo for gene therapy and other therapeutics, sgRNA and Cas9 must be efficiently delivered to target cells. This is a challenge for many methods of gene therapy. Cas9 proteins are ~1,400 amino acids, and they could be incorporated into viral vectors (e.g. adenoviral vectors and lentivirus) more easily if they were smaller [33]. It may be possible to utilize smaller cas9 by deleting domains unnecessary domains to function [33,34]. Even if this is possible, the use of viral vectors comes with its own host of problems. One of the problems is that viral vectors can cause the development of an immune response if the vector is recognized as a foreign invader. More research on the problem is necessary to resolve this complex issue [34].

Conclusion

The discovery of biological concepts can often provide a framework for the development of novel molecular tools, which can help increase understanding and manipulation life. CRISPR-Cas systems are now a useful tool kit for engineering eukaryotic cells, especially human cells.
This study discussed briefly some of the recently discovered uses of CRISPR technology, including the development of phage-resistant bacterial strains used to identify the interaction of pathogenic bacteria in a clinical setting with either phages or conjugative plasmid that can transfer antibiotic resistance and the role of CRISPR in the evolution of bacterial pathogens. The CRISPR approach is an additional tool for biologists’ use, and which can become a means for medical application. Future research on basic mechanistic details of the different CRISPR/Cas systems is expected to reveal a more complete picture of the extensive applicability of these immune complexes. The foremost problems that must be overcome are addressing CRISPR specificity and developing effective and safe delivery systems.

Conflict of interests

The authors declare that they have no competing interests.

References

11. https://www.broadinstitute.org/what-broad/areas-focus/project-spotlight/crispr-timeline

Figure 4: Major applications of the CRISPR system by genome editing in eukaryotic organisms. (A) Genome engineering with Cas9 nuclease, (B) Genome engineering by double nicking with pained Cas9 nickases, (C) Localization with defective Cas9 nuclease.


