The Development and Application of the CRISPR/CAS System as a Powerful New Tool for Genome Editing: A CASe Study

Zoe Dubrow

Biochemistry 158

1. Introduction

Only a hundred and fifty years have passed since Gregor Mendel's discovery of simple Mendelian inheritance. In a remarkably short amount of time humans have achieved such impressive feats as sequencing the entire human genome and gaining understanding of the causes of most genetic disease. However, despite the vastly increased understanding of genetics and mutation, we still struggle to specifically modify genomes at the single gene level. Genomic editing has application in a diverse range of technologies including drug development, gene therapy and synthetic biology. Despite the predicted utility of a successful gene editing technique, many current methods like Zinc Fingers Nucleases and TALENs have confounding issues like low efficiency, time-consuming procedures, and lack of specificity for both model organisms and humans. In the past several years, a new gene editing system, derived from bacteria, has arisen as a frontrunner for efficient and successful gene editing. This new system is based on clustered regularly interspaced short palindromic repeats (CRISPR). CRISPRs are DNA loci that contain multiple, short, direct repetitions of base sequences which are inherited as a result of ancestral infections. They express CRISPR RNAs which are small strands of RNA that interact with invasive DNA. Cas genes code for proteins involved in CRISPR function. 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 genetic material. Researchers can now utilize this mechanism by inserting plasmids with Cas genes and specifically designed CRISPRs into an organism. This novel technique allows us to cleave the genome of most organisms, prokaryotic or eukaryotic, at almost any site. In this paper, the origin, applications, and future challenges of the CRISPR/Cas system will be discussed.

2. Discovery of the CRISPR/Cas system

In 1987, the first description of a CRISPR locus was published by a lab which discovered a group of 29-nucleotide repeats divided by non-repetitive short sequences in E. coli, which are now known as spacers (Ishino et al., 1987). Mojica et al. (2000) later determined that similar sequences of short regularly spaced repeats (SRSRs) are present in many species of bacteria. In 2002, Jansen et al. further defined these SRSRs as a group of repeats that is preceded by both a 'leader' sequence and an AT-rich region. The authors also found a conserved family of genes that generally exist near the repeats. This family consisted of genes encoding proteins like DNA helicases (cas3) and exonucleases (cas4), which implied that Cas genes likely have function in DNA metabolism. They renamed the SRSRs as CRISPRs and the genes found near them as CRISPR-associated genes (Cas) (Jansen et al, 2002).

Bolotin et al (2005) sequenced the CRISPR spacers of 24 strains of *Streptococcus thermophilus* and *Streptococcus vestibularis* and found that the sequences were homologous with

extra-chromosomal elements. The authors theorized that CRISPR spacers are the heritable remnants of infection from exogenous DNA. Furthermore, they found that phage sensitivity is related to the number of spacers in the CRISPR locus. Therefore, the authors believe that spacers give rise to anti-sense RNA which promote an immune response against exogenous DNA. Like Jansen et al., Bolotin et al. also concludes that CRISPRs are likely involved in DNA degradation.

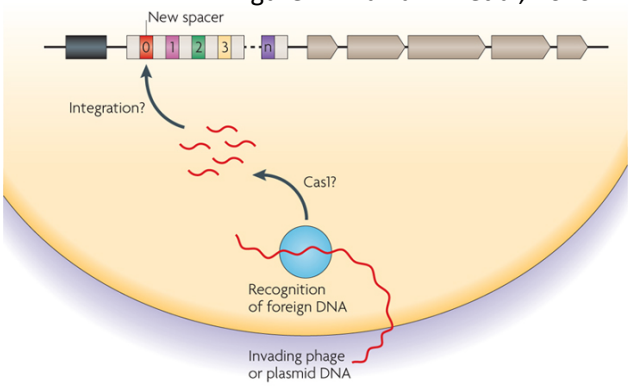
In 2006, Marakova et al. discussed the idea that CRISPR based immunity could be function through a mechanism analogous to RNAi. In 2007, Barrangou et al. determined that the sequence identity of CRISPR spacers must match sequences in a phage genome in order to induce CRISPR-based immunity. Moreover, they found that new spacers are integrated into CRISPR loci during phage infection. crRNAs recognize the "protospacer" sequences of complementary invasive plasmids. crRNAs in many bacterial species have been shown to recognize DNA targets. One experiment on *S.epidermidis CRISPRs* tested the level of CRISPR immunity when the bacteria was infected with a plasmid in which a self-splicing intron was inserted into the protospacer so that the intron was present in DNA, but spliced out of RNA (Marraffini et al, 2008). This caused CRISPR immunity to be compromised, showing that DNA is the target of the CRISPR interference system.

3. CRISPR Mechanism

There are three main types of CRISPR systems which each utilize slightly different strategies. Type II CRISPR systems will be discussed since they are most commonly used in gene editing applications. CRISPRs act in immunity through a multistep mechanism that begins with the integration of new spacers into CRISPR loci. In order to function, CRISPR RNAs (crRNAs) are then transcribed which leads to a process called CRISPR interference.

A. Spacer Integration

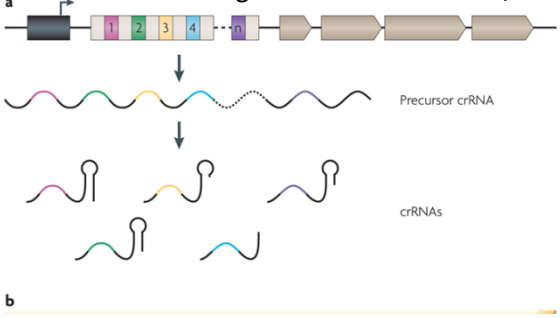
Figure 1: Marraffini et al, 2010



New spacers are inserted at the leader end of the CRISPR array when bacteria is infected with foreign DNA (Figure 1) (Marraffini et al., 2010). However, this mechanism does not cause CRISPRs to expand indefinitely; when a new spacer is inserted, a different spacer is generally deleted (Barrangou et al., 2007). New spacer integration allows for organisms to adapt immunity based on the variety of phages present in their current environment, while retaining relevant ancestral spacers. CRISPR prokaryotic adaptive immunity is heritable and based on nucleic acids which interestingly contrast eukaryotic adaptive immunity which is structured around amino acids and therefore is not inherited.

B. crRNA Transcription

Figure 2: Marraffini et al, 2010



In CRISPR Type II systems, Cas genes encode four Cas proteins, including Cas1 and Cas2 proteins

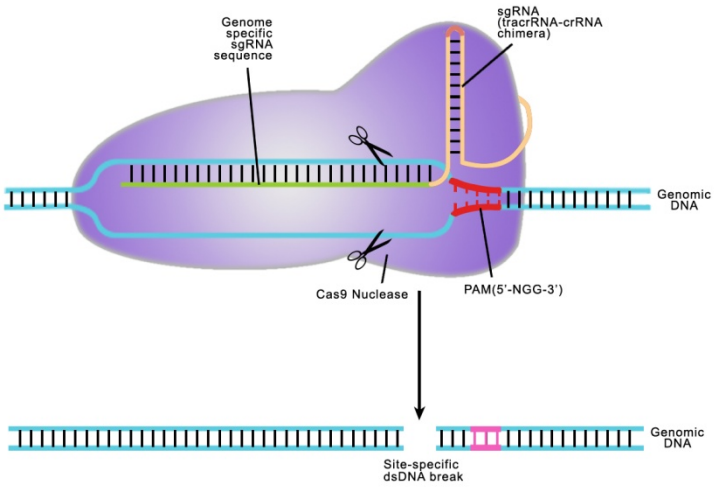
which are universal in all CRISPR systems. Type II systems also express either Cas4 or Csn2 which are involved in spacer integration. Finally, all Type II CRISPR systems encode a highly conserved Cas9 gene (Cong et al, 2013). Long primary pre-crRNA is transcribed from the CRISPR loci. Pre-crRNA pairs with trans-activating CRISPR RNA (tracrRNA) in order to be processed by RNase III (Marraffini et al, 2010). This creates shorter mature crRNAs (Figure 2). Cas9 catalyzes the formation of a crRNA-tracrRNA complex (Cong et al, 2013).

C. CRISPR Gene Editing

After transcription and processing, mature crRNAs complex with Cas9 and tracrRNA. This complex binds to a protospacer sequence of extra-chromosomal double-stranded DNA. The process is dependent on a protospacer adjacent motif (PAM) (Cong et al, 2013). When the Cas9/tracrRNA/crRNA complex binds the target sequence of the dsDNA, an R-loop forms and one DNA strand pairs with crRNA and the other disassociates. Both strands of DNA are cut near the PAM sequence. The crRNA acts as a guide while Cas9 acts as the endonuclease to cleave the DNA. The presence of double stranded breaks (DSB) in the DNA leads to activation of the DSB repair machinery either Non-homologous end joining (NHEJ) or homology-directed repair (HDR) (Cong et al, 2013). NHEJ causes insertions or deletions (indels) at the break (Mali et al, 2013) which leads to gene silencing of the invading DNA, the method through which CRISPR-mediated immunity in Type II systems functions.

4. Application of CRISPR in Eukaryotic Organisms

Figure 3: TebuBio.com



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, mouse, and even humans (Mali et al., 2013). Based on the type II CRISPR/Cas9 mechanism, researchers have engineered a RNA chimera of tracrRNA-crRNA called single guide RNA (sgRNA) which is able to cause sequence-specific binding to dsDNA (Cong et al, 2013). Cas9 nuclease is capable of interacting with this engineered sgRNA in order to specifically cleave dsDNA in regions complementary to the introduced sgRNA (figure 3). The presence of double stranded breaks (DSB) in the DNA leads to activation of the DSB repair machinery for either 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 (Mali et al, 2013).

Moreover, modified versions of this system can be used to knockout genes, insert new exogenous DNA into the host genome, and to block RNA transcription for a variety of applications.

A. Gene Knockout System

One can imagine many potential ways in which knocking out a specific gene could be useful. CRISPR could be used to delete harmful disease causing mutations in the human genome, especially in IVF embryos. Alternatively, the system could be used to knockout specific genes in model organisms in order to study diseases. The applications of CRISPR/Cas9 knockout for HIV therapy and multiple simultaneous knockout in model organisms will be discussed.

Application CRISPR Gene Knockout System as Therapy for HIV

Human immunodeficiency virus (HIV) is a lentivirus that causes acquired immunodeficiency syndrome (AIDS). HIV infection systematically destroys the human immune system. Recently, Zinc Finger Nucleases have been utilized to disrupt CCR5, a protein necessary for HIV to enter target cells (Holt et al, 2010). However, the CRISPR/Cas9 system could also be employed to combat HIV in a slightly different manner. Instead of targeting CCR5, CRISPRs can be used to disrupt the long terminal repeat (LTR) promoter of HIV (Ebina et al, 2013). This LTR is necessary for viral genomic RNA transcription. In 2013, Ebina et al. showed that disruption of the LTR region by a CRISPR system could be accomplished in HIV-1 provirus-integrated human cell lines (Ebina et al. 2013). During the course of infection, HIV integrates itself into the host genome, so while retroviral therapies can control HIV, the dormant virus still exists in the host genome. The CRISPR system has the unique potential to target the integrated genomic HIV. LTR regions exist on both sides of the integrated HIV genome. Due to this, the CRISPR system can remove the sequence of integrated HIV DNA from the host genomes by cleaving at both LTRs (Ebina et al, 2013). This appears, potentially, to be a very promising form of therapy, though it is still at a very early stage. Future challenges include determining potential off-target effects, as well as developing a safe and effective delivery system.

Application CRISPR Gene Knockout System for the Creation of Multiple Gene Knockout Model Organisms

One interesting aspect of the CRISPR/Cas9 system is that multiple sgRNAs can be inserted into cells making it possible to knock-out more than one gene (Cong et al., 2013). This is an ability unique to the CRISPR system. Wang et al (2013) describes the efficient creation of mouse stem cells with five genes disrupted at the same time using the CRISPR system. Normally, in order to create mice with multiple mutations it would be necessary to cross different mice or complete other time consuming procedures. Niu et al. (2014) describes the simultaneous disruption of two genes in single cell monkey embryos which were then inserted into surrogate mothers. This resulted in monkeys being born with mutations in the two genes targeted by the CRISPR/Cas9 system without off-target effects (Niu et al., 2014). The ability to edit the genomes of model animals, especially primates, is extremely important to medical research and will assist in the development of new treatments in future.

B. Gene Knockin System

Cong et al. studied Cas9 nickase, a variation of Cas9, which nicks target DNA in order to induce homology-directed repair (HDR) (Cong et al., 2013). Cas9 nickase can be utilized as a tool for gene insertion as well as gene knockout because the HDR pathway results in repair template substitutions rather than the deletions (or less commonly, insertions) caused by non-homologous end joining (Cong et al., 2013). This approach is particularly interesting because it decreases the probability of off-target mutagenesis (Cong et al., 2013). There are many applications for gene insertion via the CRISPR/ Cas nickase system.

Application of CRISPR/Cas Nickase Gene Knockin System in Stem Cells Therapies

Many genetic diseases could be cured by modifying genomic sequences of pluripotent stem cells of patients to express wild-type copies of the disease causing genes. This would allow for autologous stem cell therapies which reduces the risk of graft-host disease compared to allogenic treatments.

For example, Schwank et al, studied the use of CRISPR gene knock-in as a therapy for Cystic Fibrosis. They succeeded in correcting mutant CFTR Delta-F508 alleles (alleles with a mutation that causes cystic fibrosis) using the CRISPR/Cas9 mediated homologous recombination in intestinal stem cells. Schwank also showed that the corrected genes could function normally in an organoid system. Other studies have revealed that mouse organoids grown in vitro can be successfully transplanted into living mice (Yui et al., 2012). Systems like this could eventually be used for human stem cell therapy. However, this technique does give rise to the risk of endogenous gene disruption and activation of nearby oncogenes. More work must be one to accurately determine and reduce the risks of this technique.

C. CRISPR Interference (CRISPRi)

In 2013, Qi et al. created a catalytically dead Cas9 (dCas9), lacking endonuclease activity, in order to function in gene silencing as opposed to gene editing through DSBs. This method, called CRISPR interference (CRISPRi) halts mRNA synthesis by blocking RNA polymerase at the promoter region of the DNA. CRISPRi has the ability to silence multiple genes at the same time. This method can be compared to RNAi. However, CRISPRi has the possible advantage of working earlier than RNAi by silencing the gene before mRNA is created rather than simply degrading the mRNA.

Application of CRISPRi in Manipulation of Stem Cell Differentiation

CRISPRi has many possible applications in medical research. For example, Kearns et al. (2014) researched the ability of dCas9 to influence the differentiation state of human pluripotent stem cells. The authors' research determined that dCas9 could be used to positively or negatively regulate the expression of particular target genes that influence cell differentiation. CRISPRi could therefore be useful to investigate stem cell differentiation pathways (Kearns et al, 2014).

5. Alternative Genome Editing Strategies

While CRISPR is a very promising technique, it is still in its infancy. Many uncertainties still surround its mechanism and effectiveness. Importantly the exact molecular mechanism of

how sgRNAs bind to dsDNA is unknown and off-target effects of sgRNAs are likewise uncertain. Other techniques for gene editing also exist, each with their own issues.

A. Zinc Finger Nucleases (ZFNs)

Zinc Finger Nucleases were one of the first of genome editing methods to be developed. ZFNs are created by fusing a zinc finger DNA-binding domain to a DNA-cleavage domain resulting in a construct with the ability to cleave specific sequences of double-stranded DNA (Gaj et al, 2013). ZFNs are created to target unique DNA sequences in the genome of most organisms. However, unlike CRISPRS, engineering ZFNs is difficult and takes a long time to complete (Gaj et al, 2013).

B. Transcription Activator-like Effector Nucleases (TALENS)

TAL effectors (TALEs), are a genetic editing tool originally discovered in *Xanthomonas,* a pathogenic bacteria with a type III secretion system (T3S), a needle-like structure, which secretes and translocates effector proteins into plant cells. TALENs can edit the genomes of most organisms and cell types by inducing double-strand breaks, which initiates repair mechanisms in the cell to reconnect the DNA via Non-homologous end joining (NHEJ). This results in insertions or deletions (indels) in the genome. New exogenous DNA can also be introduced into a genome utilizing NHEJ. However, TALENs greatly differ in efficiency between species, cell type, nuclease, and target gene, so in some cases this method can be less predictable and effective. TALENs utilize customized TALE tandem repeats which can be difficult to construct depending on their length (Gaj et al, 2013).

While TALENs and ZFNs are clearly useful gene editing strategies, they are more difficult to engineer and apply than CRISPRs and offer few additional advantages. Overall, it seems that the CRISPR/Cas9 system may have greater potential than ZFNs and TALENs for use in many applications.

6. Future Directions of CRISPR Technology

Though CRISPR technology has come a long way in a very short amount of time, there are still many challenges that must be overcome in order for its value in applications to be fully realized. The foremost problems that must be overcome are addressing CRISPR specificity and developing effective and safe delivery systems.

CRISPR Specificity and Approaches to reduce Off-Target Interactions

Like other genome editing systems, the targeting specificity of the CRISPR/Cas9 mechanism is exceedingly important. sgRNA/Cas9 complexes have been found to be tolerant to up to 3 nucleotide mismatches with their targets (Mali et al, 2013). If this system were utilized as is, it would cause major problems in living organisms when genes other than those that are specifically being targeted were affected. Though this is certainly a cause for concern, several different methods exist to avoid these off-target effects. One potential method to increase specificity of CRISPR/Cas9 systems is requiring multiple sgRNA-Cas9 complexes for activity (Mali et al, 2013). Another way to reduce off-target effects is substituting normal Cas9 for Cas9 Nickase. Nicks in the DNA activate HDR, which leads to greatly reduced off-target effects (Mali

et al, 2013). Still, off-target effects must be further researched in order to avoid negative outcomes.

Development of CRISPR/Cas9 Delivery Systems

In order to utilize gene editing in vivo for gene therapy and other theraputics, sgRNA and Cas9 must be efficiently delivered to target cells. This is a challenge faced by 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 (Mali et al, 2013). It may be possible to utilize smaller Cas9 orthologs by deletion of domains unnecessary to function (Mali et al, 2013). Even if this is possible, the use of viral vectors comes with its own host of problems. One of these problems is that viral vectors can cause the development of an immune response if the vector is recognized as a foreign invader. More research must be completed to resolve this complex issue.

7. Conclusions

CRISPR/Cas9 is a novel technique with a bright future in genomic editing. It has the potential to be useful in a broad range of applications from simplifying research to acting as a new form of gene therapy for patients with HIV and genetic diseases. CRISPR is still a young system and more research must be completed in order to rectify its problems. While there are many challenges ahead before CRISPER/Cas9 can be utilized as a safe and reliable gene therapy, these challenges do not seem insurmountable. Research in the area of CRISPR/Cas9 is gaining speed and this system could very well be the solution to many medical issues we face today.

8. Citations

1. Bolotin, Alexander, et al. "Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extra-chromosomal origin." *Microbiology*151.8 (2005): 2551-2561.
2. Barrangou, Rodolphe, et al. "CRISPR provides acquired resistance against viruses in prokaryotes." *Science* 315.5819 (2007): 1709-1712.
3. Cong, Le, et al. "Multiplex genome engineering using CRISPR/Cas systems."*Science* 339.6121 (2013): 819-823.
4. Ebina, Hirotaka, et al. "Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus." *Scientific reports* 3 (2013).
5. Gasiunas, Giedrius, Tomas Sinkunas, and Virginijus Siksnys. "Molecular mechanisms of CRISPR-mediated microbial immunity." *Cellular and Molecular Life Sciences* 71.3 (2014): 449-465.
6. Gaj, Thomas, Charles A. Gersbach, and Carlos F. Barbas III. "ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering." *Trends in biotechnology* 31.7 (2013): 397-405.

7. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* **169**: 5429–5433
8. Jansen, Ruud, et al. "Identification of genes that are associated with DNA repeats in prokaryotes." *Molecular microbiology* 43.6 (2002): 1565-1575.
9. Haurwitz, Rachel E., et al. "Sequence-and structure-specific RNA processing by a CRISPR endonuclease." *Science* 329.5997 (2010): 1355-1358.
10. Holt, Nathalia, et al. "Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo." *Nature biotechnology* 28.8 (2010): 839-847.
11. Hou, Zhonggang, et al. "Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis." *Proceedings of the National Academy of Sciences* 110.39 (2013): 15644-15649.
12. Kearns, Nicola A., et al. "Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells." *Development* 141.1 (2014): 219-223.
13. Makarova, Kira S., et al. "A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action." *Biology direct* 1.1 (2006): 7.
14. Mali, Prashant, Kevin M. Esvelt, and George M. Church. "Cas9 as a versatile tool for engineering biology." *Nature methods* 10.10 (2013): 957-963.
15. Marraffini, L. A. & Sontheimer, E. J.CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science***322**, 1843–1845 (2008).
16. Marraffini, Luciano A., and Erik J. Sontheimer. "CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea." *Nature Reviews Genetics*11.3 (2010): 181-190.
17. Mojica, Francisco JM, et al. "Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria."*Molecular microbiology* 36.1 (2000): 244-246.
18. Niu, Yuyu, et al. "Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos." *Cell* (2014).
19. Qi, Lei S., et al. "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." *Cell* 152.5 (2013): 1173-1183.
20. Schwank, Gerald, et al. "Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients." *Cell stem cell* 13.6 (2013): 653-658.
21. Wang, Haoyi, et al. "One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering." *Cell* 153.4 (2013): 910-918.