The discovery and development of the CRISPR system in applications in genome manipulation

Journal:	<i>Biochemistry and Cell Biology</i>
Manuscript ID	bcb-2016-0159.R1
Manuscript Type:	Mini Review
Date Submitted by the Author:	25-Aug-2016
Complete List of Authors:	Lau, Veronica; University of Manitoba, Children's Hospital Research Institute of Manitoba Davie, James; University of Manitoba, Children's Hospital Research Institute of Manitoba
Keyword:	CRISPR-Cas9, genome editing

The discovery and development of the CRISPR system in applications in genome manipulation

Veronica Lau and James R. Davie

Children's Hospital Research Institute of Manitoba, University of Manitoba,
Winnipeg, Manitoba, Canada R3E 3P4

Address for Correspondence:

Dr. James R. Davie
Children's Hospital Research Institute of Manitoba
715 McDermot Avenue
Winnipeg, Manitoba, R3E 3P4
Tel. # (204)-975-7732
Fax # (204)-977-5691
E-mail: jim.davie@umanitoba.ca

Running title: History and advances of CRISPR Cas9 technology

Conflict of interest: None

Abstract

The Clustered regularly interspaced short palindromic repeat (CRISPR) CRISPR associated 9 (Cas9) system is a microbial adaptive immune system that has been recently developed for genomic engineering. From the moment the CRISPR system was discovered in *Escherichia coli*, the drive to understand the mechanism prevailed, leading to rapid advancement in the knowledge and applications of the CRISPR system. With the ability to characterize and understand the function of the Cas9 endonuclease came the ability to adapt the CRISPR-Cas9 system for use in a variety of applications and disciplines ranging from agriculture to biomedicine. This review will provide a brief overview of the discovery and development of the CRISPR-Cas9 system in applications such as genome regulation and epigenome engineering, as well as the challenges faced.

Key words: CRISPR-Cas9, genome editing

INTRODUCTION

Genome editing is a powerful tool used to manipulate the genome through insertion or deletions of DNA. The more commonly known genomic editing tools are Zinc Finger Nucleases (ZFN) and Transcription Activator-Like Effector-based Nucleases (TALEN). These methods are based on protein-DNA interactions and require enzyme engineering. In recent years, another genomic editing tool has been proposed called the Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) CRISPR associated 9 (Cas9) system. Unlike ZFN and TALEN, the CRISPR-Cas9 system uses guide RNAs to target the region of interest.

The CRISPR-Cas9 system is an adaptive immune system utilized in microbials, allowing this system to acquire resistance to foreign genetic elements derived from invading foreign DNA of phages and plasmids. The CRISPR-Cas9 system acts by integrating the nucleic acid of viruses and plasmids into the CRISPR arrays and produces short RNA sequences called CRISPR RNA (crRNA) that specifically target these invading virus genomes for degradation. This mechanism can be described in three phases: adaptation, expression and interference. There are multiple molecules involved in these phases including Cas enzymes with various functions, trans-encoding small RNAs (tracrRNAs) for processing, and a protospacer adjacent motif (PAM) for docking sites used by specific complexes. By understanding how the CRISPR Cas9 system operated with all its components, and how to manipulate the short RNA sequences to target other regions of interest, the CRISPR Cas9 system has revolutionized genomic editing in a wide variety of disciplines.

History

The first observation of the CRISPR locus occurred in 1987 when Ishino and colleagues sequenced the *iap* gene in *Escherichia coli* to identify the proteins coded by the gene (Ishino et al. 1987). In sequencing it, they discovered a peculiar pattern at the 3'-flanking region. They identified five identical sequences of 29 nucleotides separated by 32 non-repeating spacer sequences (see Figure 5 in (Ishino et al. 1987)). Soon after this discovery, these CRISPR arrays were found in other bacteria and archaea, but not in eukaryotes or viruses. To date, it has been determined that 84% of archaea and 45% of bacteria contain these arrays, with sequences of the repeat being unique to each strain (Rath et al. 2015). The term "Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)" was given to these patterned sequences by Jansen et al. in agreement with Mojica et al., who discovered the CRISPR-associated (*cas*) genes upstream of the CRISPR loci (Jansen et al. 2002; Mojica et al. 2005). Jansen and colleagues were able to identify four *cas* genes (*cas1*, *cas2*, *cas3* and *cas4*) (Jansen et al. 2002). Using database analysis, they were able to predict the function of *cas3*, which has domains similar to those of the superfamily 2 helicases and *cas4*, which has homologies to the family of RecB exonucleases.

In 2005, three research groups independently discovered that the role of CRISPR loci was in adaptive immunity (Bolotin et al. 2005; Mojica et al. 2005; Pourcel et al. 2005). They noticed that the spacer sequences were not random; rather, they seemed to be derived from phage DNA and conjugative plasmids. To further understand the relationship between CRISPR loci and extrachromosomal sequences, CRISPR profiles were compared to phage sensitivity in *S. thermophiles* (Bolotin et al. 2005). A negative correlation was observed between the two factors strengthening the speculation that the CRISPR loci were related to phage sensitivity. A lower number of repeat spacer pairs correlated with a higher phage

sensitivity. The hypothesis that CRISPR was involved in adaptive immunity was further validated by Barrangou et al. who manipulated the spacer sequences in the CRISPR loci in phage resistant *S. thermophiles* (Barrangou et al. 2007). When the spacer sequences were deleted, the phage resistant strains became sensitive to the phage, confirming that CRISPR loci were involved in adaptive immunity. Bolotin et al. proposed that CRISPR loci used antisense RNA to inhibit phage gene expression (Bolotin et al. 2005) and Mojica and colleagues believed that the CRISPR system functioned similarly to eukaryotic small interference RNA (Mojica et al. 2005). It was discovered later that the CRISPR Cas systems integrated small foreign DNA into the CRISPR locus. Small RNA profiling studies suggested that CRISPR spacers and repeat sequences were transcribed as a long transcript and processed into small CRISPR related RNAs called crRNAs (Brouns et al. 2008; Marraffini and Sontheimer 2010).

The discovery that CRISPR repeats formed an RNA secondary structure aided in the division of the different CRISPR subtypes. The repeat sequences could either form a structured hairpin RNA or an unstructured RNA (Kunin et al. 2007). In 2011, tracrRNAs were discovered through the differential RNA sequencing of *Streptococcus pyogenes* (Deltcheva et al. 2011). These tracrRNA displayed almost exact complementation to the crRNA with one mismatch. It was determined that the processing site of both tracrRNA and pre-crRNA occurred in RNA duplex regions and that tracrRNA was necessary for maturation of crRNA. Since none of the Cas proteins contain RNase III-like motifs, RNase III of the organism is recruited to the complex and cleaves the RNA dimer of pre-crRNA and tracrRNA. Of all the Cas proteins, only Csn1 is required for the maturation of crRNA and tracrRNA cleavage in *S.pyogenes*. Deltcheva and colleagues proposed that the role of Csn1 in the maturation of

crRNA was to facilitate base-pairing between tracrRNA and crRNA for RNase III recognition and cleavage. Csn1 was also given the name Cas9 (see Figure 4 in (Deltcheva et al. 2011)).

By understanding the concept of how Cas9 worked, Jinek et al. were able to design a remarkable technique for genomic editing (Jinek et al. 2012). Their study focused on the mechanism of Cas9 in targeting and cleaving foreign DNA. For the CRISPR-Cas system to work for genomic editing, some key components were necessary. They determined that Cas9, mature crRNA and a PAM were not sufficient for Cas9-catalyzed plasmid DNA cleavage. However, with the addition of tracrRNA, a crRNA that binds the target DNA (also called protospacers) and magnesium, the cleavage reaction was successful. Jinek and colleagues did multiple tests to determine the role of Cas9 and tracrRNA in CRISPR-Cas (Jinek et al. 2012). Once they understood this mechanism, they created small guide RNAs (sgRNA) which mimicked the tracrRNA-crRNA complex and were able to reduce the requirements for a programmable system. With the RNA-programmed Cas9, any DNA sequence could be targeted and cleaved.

After Jinek tested the proposition that the CRISPR-Cas9 system could be used for genomic engineering, Cong et al. were able to engineer Cas9 nucleases to generate precise DNA cleavage in human and mouse cells (Cong et al. 2013). Cong and colleagues transfected 293FT cells with the CRISPR-Cas9 derived from *S. pyogenes* and determined that the CRISPR Cas9 system could function in these cells with the addition of nuclear localization signatures. Furthermore, they showed that the CRISPR-Cas9 system could be used to facilitate multiplexed genome engineering. By using a single CRISPR array, they were able to successfully cleave both the EMX1 and PVALB by encoding an EMX1- and PVALB-protospacer. The CRISPR-Cas9 system was further manipulated and repurposed for other genomic applications such as genomic regulation (Qi et al. 2013).

Mechanism of CRISPR in microbial defense

The molecular mechanism of the CRISPR mediated immunity can be described in three phases; the adaptive phase, the expression phase and the interference phase (Deltcheva et al. 2011; Jinek et al. 2012; Makarova et al. 2011; Marraffini and Sontheimer 2010; Rath et al. 2015) (Figure 1). During the adaptive phase, new spacers derived from plasmids and phages are incorporated into the CRISPR loci (Marraffini and Sontheimer 2010). This allows the cell to adapt to the invading plasmid or phage in the environment and inactivate invaders. The new spacers are inserted adjacent to the AT rich leader sequence that is the intervening sequence between the Cas genes and the CRISPR locus. The CRISPR arrays are limited in size, thus to incorporate new spacers previously conserved spacers must be removed. The process of spacer acquisition can be divided into two steps: i) selection of protospacer and generation of spacer material and ii) spacer integration into CRISPR loci. The mechanism of how spacers are integrated is not completely understood; however, Cas1 and Cas2 proteins have been shown to have a key role in this process. The importance of Cas1 and Cas2 were determined when overexpression of these two proteins in *E. coli* were sufficient for spacer integration to occur even in the absence of the other Cas proteins. If the active site of Cas1 is mutated, spacer integration in *E. coli* is abolished. Cas1, a metal-dependent DNase and Cas2, a metal-dependent endoribonuclease, form heterodimers. Other molecules that may have a role in spacer integration are Cas proteins Cas9 and Csn2 and tracrRNA. However, the role of Csn2 and tracrRNA in spacer acquisition is unknown, whereas Cas9 may play a role in guiding the integration machinery to sequences that act as recognition signals for endonucleases to cleave protospacers (Rath et al. 2015).

During the expression phase, short crRNAs complementary to the protospacers are created to trigger degradation of target sequences. For this to occur, CRISPR transcription

must be initiated at the end of the locus containing the leader sequences. The repeats and spacers are transcribed as a long transcript and processed into short crRNAs with the aid of tracrRNA. This process is catalyzed by endoribonucleases which are either part of a larger subunit (such as CRISPR associated complex for antiviral defence (Cascade) or a single enzyme) (Marraffini and Sontheimer 2010).

In the interference phase, the invading DNA or RNA is targeted and cleaved. The crRNA acts as a guide and recruits Cas protein complexes to the complementary viral or plasmid derived target sequences. During this step, short conserved regions adjacent to the protospacers, called PAM, may act as a recognition site for complex docking. In most strains, such as *S. thermophiles*, *E. coli* and *Staphylococcus epidermidis*, the crRNA targets the foreign DNA. However, there are some strains such as *P. furiosus* where crRNA targets foreign RNA instead. Multiple Cas proteins are involved in one or more stages of the CRISPR-Cas system (Marraffini and Sontheimer 2010; Rath et al. 2015).

Classification of CRISPR-cas system

The original classification of the CRISPR classification consisted of 8 different subtypes that were categorized by genomic context (Haft et al. 2005). These included subtypes Ecoli, Ypest, Nmeni, Dvulg, Tneap, Hmari, Apern, and Mtube. Each subtype consisted of a combination of core proteins (cas1-6) which are characterized by their location in relation to the CRISPR loci. However, this classification was revised, as it did not consider the evolutionary relationship of the CRISPR system. A new classification system was then proposed which was classified based on presence or absence of certain core Cas proteins (Makarova et al. 2011; Makarova et al. 2015)(Figure 1).

The new classification consisted of three types of CRISPR-Cas systems; type I, II and III. Type I is defined by the presence of the Cas protein Cas3 which has both helicase and DNase domains for target sequence degradation. There are six different subtypes of type I systems which vary in number of *cas* genes. All type I systems encode a Cascade-like complex which contains the Cas3 protein, Cas 1 and Cas2. The core enzyme identified in the Cascade complex is a Repeat-Associated Mysterious Protein (RAMP) with RNA endonuclease activity, catalyzing the processing of pre-crRNA transcript (Makarova et al. 2011). The RAMP superfamily contains modular genes that are always associated with one another in the CRISPR locus. RAMPs form a large superfamily of Cas proteins. They are characterized by a glycine-rich loop, and they contain at least one RNA recognition motif. Some RAMPs may contain sequence or structure specific RNase activity which is involved in processing pre-crRNA transcripts (Marraffini and Sontheimer 2010; Rath et al. 2015).

Type II systems uses the Cas9 protein as well as Cas1 and Cas2. The Cas9 is a large endonuclease that is essential in crRNA processing and cleaves target DNA. Thus in type II systems Cascade complexes are absent. Cas9 consists of at least two nuclease domains, a HNH (His-Asn-His domain) and a RuvC nuclease domain and has a role in cleaving target sequences producing blunt ends. In type II systems, tracrRNA hybridizes with the nearly complementary region in pre-crRNA during the production of crRNA (Deltcheva et al. 2011). Cleavage of the duplex is carried out by RNase III in the presence of Cas9, which is thought to be necessary for crRNA-guided silencing of targeted DNA (Qi et al. 2013). There are three subtypes of type II. Type II-A encodes an addition protein Csn2, type II-B encodes Cas4, and type II-C lacks an additional protein (Rath et al. 2015). Type II is only present in bacterial systems, where type II-C is the most common subtype present in bacteria (Makarova et al. 2015; Rath et al. 2015).

The final type, Type III, contains protein Cas10. The function of this protein, however, remains unclear. Depending on the subtype of type III systems, Cas proteins can be either in the Csm (in type III-A) or Cmr (in type III-B) complexes which are similar to Cascade. Like type I and II systems, type III systems can target DNA; however, in addition to DNA, type III systems can also target RNA. In type III systems, RAMP seems to be involved in the processing of spacers. Type I and type III systems are present in both bacteria and archaea.

A recent study found evidence for two new putative types; type IV and type V (Makarova et al. 2015). Type IV is often found on plasmids in bacterial genomes and is the most similar to type III-B. Type IV systems are usually located away from the CRISPR arrays, or in genomes with CRISPR arrays that have not been detected, and may be mobile modules that use available crRNA from different CRISPR arrays. Type IV systems also lack the Cas genes *cas1* and *cas2*. Instead, type IV systems use a multisubunit crRNA-effector complex comprised of a large and sometimes a small subunit composed of the proteins Cas5, Cas7 and Csf1. The *csf1* gene is the main characteristic of type IV systems. Type IV systems are divided into two variants based on the presence and absence of the family helicase DinG.

The putative type V is characterized by the presence of the Cpf1 interference module and an adaptor module. This putative type has been found in many bacterial genomes, and at least one archaeal genome. This CRISPR type contains the *cas1* and *cas2* gene and is adjacent to CRISPR arrays. The Cpf1 protein is similar to the Cas9 protein and contains a similar RuvC-like nuclease domain as well as similar Arg-rich regions. The protein also contains the TnpB transposon and the TnpB zinc finger, but lacks the HNH nuclease domain, which is present in Cas9 proteins. Type V systems are the most similar to type II subtypes. It is most similar to type II-B if the type V locus contains the *cas4* gene; however, if the locus lacks the *cas4* gene then the architecture is similar to type II-C (Makarova et al. 2015).

Utilizing the CRISPR Cas system for genome editing

Genome editing creates much potential for understanding the genome across different fields such as medicine, biotechnology and basic science. The principle for genomic editing can be summarized in two steps (Cong et al. 2013; Hsu et al. 2014; Laufer and Singh 2015; Ramalingam et al. 2013; Ran et al. 2013a; Ran et al. 2013b; Rath et al. 2015; van Erp et al. 2015). The first is introducing DNA cleavage into the targeted area, and the second step is the natural process of DNA repair. DNA can be repaired by the nonhomologous end joining (NHEJ) pathway, or the homology directed repair (HDR) pathway. The NHEJ pathway can be used to mediate gene knockouts by causing frameshift mutations and premature stop codons. HDR can be used to introduce specific donor sequences with the use of a template. ZFNs and TALENs are genome editing techniques that were previously discovered (Hsu et al. 2014; Laufer and Singh 2015; Makarova et al. 2011; Makarova et al. 2015; Ramalingam et al. 2013; Ran et al. 2013a; Ran et al. 2013b; Rath et al. 2015; van Erp et al. 2015). ZFNs were discovered in 1996 by Kim et al. (Kim et al. 1996; Ramalingam et al. 2013). This application uses zinc finger motifs which contain a $\beta\beta\alpha$ structure with a unique 30 amino acid sequence that could be modified for targeting specific DNA sequences (Huisman et al. 2015) (Chen et al. 2014). The sequence would then be cleaved using the non-specific cleavage domain of FokI endonuclease with zinc finger proteins. Each zinc finger motif binds to DNA by inserting the α helix into the major groove and can recognize 3 to 4 bp. The concept of the genomic editing technique TALEN is similar to ZFN; however, instead of using zinc finger motifs, TALENs uses TAL effectors for DNA recognition followed by cleavage by the FokI endonuclease domain. The TAL effectors can recognize a single base pair and the recognition specificity relies on repeat variable di-residues (RVDs). RVDs are repeats that are highly conserved, with high variability at amino acid positions 12 and 13. Unlike ZNFs, when

binding to DNA sequences TAL effectors do not interfere with the DNA structure, making it ideal for designing custom artificial DNA nucleases.

The CRISPR-Cas9 system differs from both of these genomic editing techniques. It uses sgRNA and does not require enzyme engineering. The protocol for the CRISPR-Cas9 system has been developed for genomic editing (Ran et al. 2013b) (Figure 2). The first step is selecting an appropriate sgRNA that is specific to the target of interest. An online CRISPR design tool (<http://tools.genome-engineering.org>) was created to find suitable target sites for the sgRNAs. The program provides sequences that are required for making sgRNA constructs, while considering the efficiency of the target modification, the potential off-target cleavage sites and the existence of a PAM sequence (Ran et al. 2013a). Li et al. have designed a statistical algorithm called Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout (MAGeCK) which effectively identifies more specific sgRNAs in addition to genes and pathways that are potential targets for CRISPR-Cas9 knockout screens (Li et al. 2014). The MAGeCK algorithm normalizes the read counts for different samples by adjusting for library size and distribution of read counts. It then uses a robust ranking aggregation (RRA) algorithm which is able to identify which pathways are more likely to be targeted by sgRNA by looking at the frequency of sgRNA specificity being constantly higher than the control (Li et al. 2014).

Once the sgRNA sequences are selected, the sgRNA must be constructed and delivered into the cell of interest. There are two methods that can be used to deliver sgRNA into a cell; it can be delivered either as a PCR amplicon containing expression cassettes, or as a sgRNA-expression plasmid (Ran et al. 2013b). A PCR amplicon is prepared by adding the custom sgRNA sequence onto the reverse PCR primer that amplifies the U6 promoter template. The amplicon can then be co-transfected with a Cas9 expression plasmid. To

construct a sgRNA-expression plasmid, a single cloning step is done. A partially complementary oligonucleotide pair, which encodes the guide sequence, is annealed and ligated into a Cas9 containing plasmid. The plasmid also contains the remainder of the sgRNA as an invariant scaffold adjacent to the oligo cloning sites. These expression plasmids can be modified for *in vivo* delivery by enabling virus production. However, a few additional constructs are required. In addition to the expression plasmid, a backbone construct, a Cas9 endonuclease, a marker selection, such as 2A-GFP and 2A-Puro fusion constructs and an antibiotic resistance vector are needed. The Cas9 endonuclease guided by the sgRNA creates double stranded nicks. When this occurs, the sequences are repaired through NHEJ or HDR. The system can tolerate certain DNA mismatches and this can promote off-target mutagenesis. Substitution of the amino acid at position D10A converts Cas9 nuclease catalytic activity to a nickase enzyme, which makes single strand breaks. This nickase will favour homologous recombination at targeted sites and reduce off-target products. Ran et al. have also developed a method that uses the D10A mutant of Cas9 plasmid in addition to a pair of sgRNA. This combination produces DNA double strand breaks with higher specific genome editing (Ran et al. 2013a).

There have been many different approaches developed to introduce the CRISPR-Cas9 system into an organism. The traditional approach of transfecting the CRISPR-Cas9 system through expression plasmids can be challenging and has certain limitations. Since these plasmids are quite large in size, they can be difficult to transfect into the cell and may be toxic to some cell lines (Kim et al. 2014). Wang et al. were able to microinject Cas9 mRNA and gRNA into mouse embryonic stem cells, resulting in 95% biallelic mutation in newborn mice (Wang et al. 2013). With this approach T7 promoter is used to recruit the T7 RNA polymerase instead of the U6 promoter. To make the process of integrating the CRISPR-

Cas9 system more rapid, Cas9 proteins/gRNA ribonucleoprotein (RNP) complexes were directly introduced into the cell, and thus removing the need for cellular transcription and translation of Cas9. The method has been successfully used to create gene-knockout mice with a mutant frequency as high as 93% in newborn mice (Sung et al. 2014). For multiplex genome editing, lentiviral delivery of a single vector with expression for Cas9 and multiple sgRNAs can be used (Liang et al. 2015b).

If the desired product is a point mutation, repair templates must be designed after the sgRNA has been delivered into the cell. This method uses HDR repair to create the specific point mutation based on the repair templates. Repair templates can be designed either as plasmid-based donor repair templates, or as single stranded oligodeoxyribonucleotides (ssODNs). Both require areas of homology with the region flanking the site of repair and can be used for the introduction of mutations of a single base pair or larger (Ran et al. 2013b).

The next step is to transfect the cells with sgRNA in the form of either the PCR amplicon expression cassette or expression plasmid. Next is the clonal isolation of cell lines, which can be done through serial dilution or Fluorescence Activated Cell Sorting (FACS). Once the clone has been isolated, an expansion period is required to allow for propagation of the isolated clonal cell line (Ran et al. 2013b).

The final step in this process is functional testing. Since NHEJ usually produces insertion and deletion mutations, to detect if the transfection was successful in incorporating the right plasmid, or PCR amplicon, sequencing or the use of the SURVEYOR nuclease assay can be done. The SURVEYOR nuclease assay detects mismatch repairs and cleaves them. The product can then be analyzed using gel electrophoresis. If the HDR pathway was used, the product can be detected using PCR amplification followed either by sequencing or restriction-fragment length polymorphism (RFLP) assays (Ran et al. 2013b).

CRISPR-Cas system for genomic regulation

The CRISPR-Cas9 system has traditionally been used, as described above, to create mutations in the genome. The highly specific targeting of Cas9 facilitated by the sgRNA lends itself to other functions as well that need not rely on the introduction of mutations. In genome regulation applications, a catalytically dead Cas9 protein is used. It was discovered that coexpression of the Cas9 lacking endonuclease activity and gRNA created a DNA recognition complex, and this complex functioned in interfering with transcriptional elongation. This technique was termed CRISPR interference (CRISPRi). With the use of the native elongating transcript sequencing (NET-seq) approach, global profiles of the position of the elongating form of the RNA polymerase was detected in *E. coli*. It was determined that the Cas9:sgRNA complex physically collided with the elongating RNAP and interfered with transcription (see Figure 1 in (Qi et al. 2013)). The CRISPRi displayed moderate repression in mammalian cells, with efficiency of repression appearing dependent upon the targeting locus (Qi et al. 2013).

CRISPR-Cas systems and epigenome studies

The CRISPR-Cas9 system can also be used in modifying and understanding the epigenome. The study of epigenetics can be done by genome editing or by epigenome editing (Chen et al. 2014; Huisman et al. 2015; Ran et al. 2013b). These forms of genetic engineering have different approaches in modifying epigenetic marks. Genome editing as stated previously, uses a nuclease to modify the target DNA sequence. Epigenome editing however, does not modify the DNA sequence. Instead of using the nuclease activity, an effector domain is used. This is done by using a modified Cas9 protein that is catalytically inactivated with effector domain conjugated to the Cas9. The domain is a small portion of a regulatory protein of interests and is utilized to create an epigenetic mark that alters transcription (see Figure 1 in (Laufer and Singh 2015)). These effector domains can either be

used to activate transcription or repress it. These domains are derived from viral elements. The original domain used to activate transcription was the VP16 domain and was derived from the Herpes Simplex Viral protein 16. This domain was then engineered as the VP64 domain, which is a fusion of four tandem copies of the VP16 domain. VP64 domain functions by recruiting the acetyltransferase p300, which acetylates histone H3 lysine 27 (H3K27ac), an epigenetic mark that is associated with active transcription. The repressive domain is usually derived from Krüppel-associated boxes (KRAB). KRAB is a potent repressor domain that is naturally occurring and is used by many zinc fingers proteins or TFs. It functions by recruiting histone modifying domains which decrease transcription activating mark histone H3 lysine 4 trimethylation (H3K4me3) and increases repressor marks such as histone H3 lysine 9 trimethylation (H3K9me3) and histone H3 lysine 27 dimethylation (H3K27me2) (Ran et al. 2013b).

Genome editing in contrast, can be used to modify the epigenome by editing proteins involved in maintaining the epigenome. The CRISPR-Cas9 system has been used to knockout the insulator CCCTC-binding factor (CTCF) in Hox clusters (Narendra et al. 2015). This resulted in euchromatin expansion into heterochromatin domains and disrupting topological boundaries. Another study has used the CRISPR-Cas9 system to delete the protein chromodomain helicase DNA-binding protein 5 (Chd5) in mouse embryonic stem cells (Hayashi et al. 2016). This protein is a chromatin remodeler that recognizes unmodified H3 lysine 4 (H3K4un) and H3K27me3 and is involved in spermatogenesis and neuronal differentiation. Chd5 has also been shown to be a tumor suppressor. From their experiment, Hayashi et al. determined that Chd5 repressed expression of the retrotransposon murine endogenous retrovirus-L (MuERV-L/MERVL) in mouse embryonic stem cells by decreasing

histone H3 variants H3.1 and H3.2 and increasing H3K27me3 mark around MuERV-L/MERV1 (Hayashi et al. 2016).

CRISPR Cas system in disease treatment

With the ease of design and the flexibility in targeting DNA sequences, the CRISPR Cas9 system has been used in the treatment of disease. Further, the development of new animal models as well as cellular models for disease studies has greatly benefitted from the application of the CRISPR-Cas9 system. Three independent studies have successfully removed the mutation in the dystrophin protein responsible for the most common form of Duchenne's muscular dystrophy (Dmd). By removing the mutated Dmd exon 23 of the gene via adeno-associated virus-9 (AAV9), and produced a shorter version of the protein which resulted in a partial rescue of muscle function (Calos 2016; Long et al. 2016; Nelson et al. 2016; Tabebordbar et al. 2016). The CRISPR Cas9 system has also been used in correcting cataracts in mice by injecting zygotes with the CRISPR Cas9 system with the specific gRNA for the mutant *CRYGC* gene allele, which encodes a member of the beta/gamma-crystallin family of proteins (Wu et al. 2015; Yang et al. 2016). Firth et al. were able to successfully generate induced pluripotent stem cells (iPSCs) from cystic fibrosis patients with the corrected F508 deletion in the cystic fibrosis transmembrane regulator (CFTR) gene by using the CRISPR Cas9 system (Firth et al. 2015). When these corrected cells were induced with chloride channel currents, the CFTR phenotype was rescued (Firth et al. 2015). Pig models for Parkinson's disease were generated via the CRISPR Cas9 system by targeting genes for DJ1, Parkin and Pink1 (Yang et al. 2016). CRISPR Cas 9 was applied to knock out the *p53* gene to generate *p53* biallelic mutant monkeys (Chen et al. 2016). Targeting the cleavage of the β -globin gene (*HBB*) in human tripronuclear zygotes was accomplished using the CRISPR Cas 9 system (Liang et al. 2015a). In their study 28 of the 54 embryos had the *HBB* gene

were successfully cleaved by Cas9, providing an efficiency of 54%. However, off target cleavage at the *OPCML* intron and *TULP1* intron were detected in the embryos, but not in the HEK293T cells (Liang et al. 2015a). This study highlighted the need to improve the fidelity and specificity of the CRISPR Cas9 approach before attempting clinical applications.

CRISPR Cas system for industrial use

CRISPR has revolutionized many fields of science. The ability to do rapid and efficient genomic editing and genomic regulation has resulted in CRISPR applications in a range of industries (see Table 1 in (van Erp et al. 2015)). The first application of CRISPR for commercial use was to strengthen viral immunity in bacterial cultures for the production of cheese and yogurt. This was initially pioneered by the company Danisco (DuPont) (Hsu et al. 2014; van Erp et al. 2015). The CRISPR technique has also expanded into the field of agriculture. Dow Agrosience and Sangamo Biosciences have collaborated to develop intellectual property with the aim of using Cas9 to genetically modify crop. Recombinetics Inc. uses the CRISPR-Cas9 system as well to improve the productivity of the livestock industry (van Erp et al. 2015). Another application that has utilized CRISPR is controlling invasive species (Webber et al. 2015). CRISPR Cas9 introduces a deleterious trait such as reduced fertility and chemical sensitivity into the population of invasive species, causing rapid reduction in the targeted population, making it more controllable. There are however, others that believe that this technique is dangerous and there are many potential risk involved with the use of CRISPR-Cas9 in the control of invasive species. One major risk is the potential of off-target mutations that could be introduced into the gene pool and potentially transferred to non-target organism. Another risk is the potential of unexpected effects that could create cascades that could be more of a threat than the target species itself (Webber et al. 2015). The CRISPR technology has also been used in development of vaccines. The vaccinia virus

(VACV) vectors have been used for vaccination and were used in the eradication of smallpox. The application of the CRISPR-Cas9 system has improved the efficiency of marker-free VACV vectors (Yuan et al. 2015). Others have used the CRISPR-Cas9 system to target specific viruses such as hepatitis B. The CRISPR-Cas9 system was used to target and cleave regions of the hepatitis B genome, preventing viral gene expression and replication (Ramanan et al. 2015).

Strengths and limitations of the CRISPR-CAS9 system

The CRISPR-Cas9 system has shown considerable potential as a leading genome editing technique. Compared to the techniques previously discovered such as ZFN and TALEN, which use DNA-protein interactions, CRISPR-Cas9 uses Watson crick base pairing between the DNA and sgRNA. When doing genomic editing with DNA-protein interaction techniques, new proteins must be engineered for each new target (van Erp et al. 2015). However, with the use of Cas9, it is not required to produce new proteins, only new sgRNAs, making this technique more rapid and easier to design. The diversity of PAM sequences for Cas protein binding between different organisms provides a choice of CRISPR-Cas systems and allows for a higher chance of targeting the sequence of interest. Because of this vast diversity, in one experiment, 20,000 different genes in the human genome and 1800 microRNAs were generated using Cas9 targeting sgRNAs. Experiments done by Cho et al. determined the efficiency of Cas9 cleavage activity (Cho et al. 2013). From their experiments, they found that the frequency of knockouts or generation of mutations ranged from 2%-18% which was similar to the efficiency of ZFN and TALENs. They also tested the specificity of the Cas9 enzyme by subjecting it to a plasmid with sgRNA and a control plasmid which lacked the gRNA. Cas9 was able to selectively cleave the plasmids with the sgRNAs and not the control plasmids. With the ease of design, efficiency and specificity, the CRISPR-Cas9

system is optimal for high throughput and multiplexed gene editing and has allowed genomic editing to be possible in the majority of model organisms, which was something that could not be done previously (Zhang 2015).

There are some unfortunate limitations to this system. As stated previously, for the CRISPR-Cas9 system to work for genomic editing, it requires the Cas9 protein, the sgRNAs and the PAM sequence. Thus each of these components must be optimized for the condition it is presented in. Since the Cas9 requires a PAM sequence, if there is no suitable PAM sequence, then the DNA cannot be targeted by Cas9. Another limitation to this technique is off target products. Although this issue has been minimized by modification of Cas9 endonuclease activity and the addition of sgRNAs, the flexibility of Cas9 binding and the characteristic of the target sequence may still lead to off target effects. This can be quite an issue when specifically designing sgRNAs to target an isoform of a gene. If the alternative splice site is further downstream from the 5' end, it may be difficult to induce a frameshift mutation that would completely knockout that specific isoform gene. There is also the issue of potentially affecting both isoforms and being unable to determine specifically which isoform is being knocked down. When attempting to work with HDR regulated genome editing, there is limitation in the mutation frequency. Since the occurrence of HDR is rare, it can be difficult to isolate, especially in cell lines. The cell lines used is also something one must consider. If the cell line of interest is deficient in a certain repair system, it will limit the types of utilizable genomic editing. Furthermore, the experiments done by others have shown that introducing the Cas9 system into certain cell lines can cause a change in cell phenotype and may be cytotoxic (Laufer and Singh 2015).

Ethics

With the advancements made for utilizing CRISPR-Cas9 system as a genome editing tool, many concerns have been raised, and have sparked the debate over the morality of the use CRISPR-Cas9 system in human embryos as well as animal model. Due to the ease of genome editing as well as the efficiency of the CRISPR-Cas9 system, it has potential to advance studies on human disease, and even cure many genetic diseases. However, should these tools be used on humans? The issue that is continuously debated with the use of the CRISPR-Cas9 tool for human genome editing is determining the boundaries of the use of this technique. Many factors in implicating the CRISPR-Cas9 system in human gene therapy that have been considered are: 1) the fear of the initial application of the CRISPR-Cas9 for disease treatment could be expanded to human germline genome modification with questionable ethical reason for doing such and 2) the unforeseen consequences that could lead to potential heritable mutations within the germline (Baltimore et al. 2015; Caplan et al. 2015).

The publication of using the CRISPR-Cas9 system in 3NZ human embryos (Liang et al. 2015a) sparked a lot of controversy and criticism from scientists in Western countries (Zhai et al. 2016). Many criticized the research done in China to be premature and unsafe. They further went and stated that the research conducted in China was overstepping ethical boundaries that have been established in Western countries (Zhai et al. 2016). However, to date, no consensus has been reached for the international guidelines on the use of CRISPR-Cas9 in human genome editing (Lander 2015; Zhai et al. 2016). At least four key issues should be considered when developing a framework to evaluate genome editing in human embryos; 1) If the editing can be done sufficiently and precisely with minimal risks, 2) if the medical reason of use is worth the potential risks involved with the editing, 3) which

individuals would have to right to agree to the genome editing, and 4) the morality involved with using CRISPR-Cas9 genome editing and its influence on society. Some suggest that further research is required before the CRISPR-Cas9 system is used on human embryos (Lander 2015). However, the UK Human Fertilisation and Embryology Authority has recently approved the work of Kathy Niakan at the Francis Crick Institute in London on using CRISPR-Cas9 on human embryos, with the conditions that the embryos are destroyed seven days after experiments (Callaway 2016).

Although human genome editing with CRISPR-Cas9 is a prevalent concern, another issue in using CRISPR-Cas9 for genome editing is the creation of genetically modified organisms (GMO) and the lack in proper regulations as well as the potential effects these GMO may have on the environment as well as for human consumption (Caplan et al. 2015). In the US, regulation for the use of genetically modified animals is done by multiple agencies and includes; the Coordinated Framework for the Regulation of Biotechnology, Food and Drug Administration (FDA), US Department of Agriculture (USDA) and the Environmental Protection Agency (EPA). Many other countries that are involved in using genetically modified animals also have regulatory systems. However, with the exception of Biological and Chemical Weapons Convention, there is no unified standard international guidance for the modification of organisms that are non-human (Caplan et al. 2015). The use of the CRISPR-Cas9 has also been considered for the implication in human food consumption. Also in this context, one is reminded of the unwanted traits in breeding poultry, resulting in an aggressive broiler breeder male, a “rapist rooster” (Grandin and Johnson 2009; Millman et al. 2000). The “rapist rooster” was a product of the broiler industry to produce chickens with more breast meat. The breeding program generated roosters that lost the art of courtship which is part of the mating protocol. The disinterested hen frustrated the hyperaggressive rooster which

violently attacked, sometimes killing, the hen (Grandin and Johnson 2009; Millman et al. 2000). With the use of CRISPR-Cas9, cattle and other food stocks can be manipulated to produce better quality food through genetic modifications that would be more favourable for human consumption. However, little is known about the effects these modified organisms would have on humans when consumed. Another concern involving genetically modified non-human organisms is the eradication of invasive species. As mentioned previously, the major ethical concerns in attempting to control invasive species is the potential of hereditary genetic mutations and the unknown effects it may have on the ecosystems (Caplan et al. 2015; Hayashi et al. 2016).

Conclusion

The progress in understanding the CRISPR system as an adaptive immunity in bacteria and archaea to utilizing the potential of the CRISPR-Cas9 system for genomic engineering has been rapid and has advanced knowledge in many different fields. With the ease of design, efficiency and specificity, the CRISPR-Cas9 system has been manipulated for different purposes including epigenome editing and genomic regulation. However, the application of CRISPR-Cas9 system in genetic engineering has brought about many ethical concerns. Although there are still challenges with the current CRISPR-Cas9 system, the rapid development of this technology will contribute to the advancement of science and the translation of this science in industry and health.

Acknowledgements

This work was supported by a Research Manitoba Studentship in partnership with CancerCare Manitoba Foundation and a Canada Research Chair to JRD. The authors

acknowledge the strong support of the Children's Hospital Research Institute of Manitoba and CancerCare Manitoba Foundation.

Draft

References

- Baltimore, D., Berg, P., Botchan, M., Carroll, D., Charo, R.A., Church, G., et al. 2015. Biotechnology. A prudent path forward for genomic engineering and germline gene modification. *Science* **348**: 36-38. doi: 10.1126/science.aab1028.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709-1712. doi: 10.1126/science.1138140.
- Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**: 2551-2561. doi: 10.1099/mic.0.28048-0.
- Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., et al. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**: 960-964. doi: 10.1126/science.1159689.
- Callaway, E. 2016. UK scientists gain licence to edit genes in human embryos. *Nature* **530**: 18. doi: 10.1038/nature.2016.19270.
- Calos, M.P. 2016. The CRISPR Way to Think about Duchenne's. *N Engl J Med* **374**: 1684-1686. doi: 10.1056/NEJMcibr1601383.
- Caplan, A.L., Parent, B., Shen, M., and Plunkett, C. 2015. No time to waste--the ethical challenges created by CRISPR: CRISPR/Cas, being an efficient, simple, and cheap technology to edit the genome of any organism, raises many ethical and regulatory issues beyond the use to manipulate human germ line cells. *EMBO Rep* **16**: 1421-1426. doi: 10.15252/embr.201541337.
- Chen, H., Kazemier, H.G., de Groote, M.L., Ruiters, M.H., Xu, G.L., and Rots, M.G. 2014. Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. *Nucleic Acids Res* **42**: 1563-1574. doi: 10.1093/nar/gkt1019.
- Chen, Y., Niu, Y., and Ji, W. 2016. Genome editing in nonhuman primates: approach to generating human disease models. *J Intern Med*. doi: 10.1111/joim.12469.
- Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**: 230-232. doi: 10.1038/nbt.2507.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819-823. doi: 10.1126/science.1231143.
- Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., et al. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**: 602-607. doi: 10.1038/nature09886.
- Firth, A.L., Menon, T., Parker, G.S., Qualls, S.J., Lewis, B.M., Ke, E., et al. 2015. Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. *Cell Rep* **12**: 1385-1390. doi: 10.1016/j.celrep.2015.07.062.
- Grandin, T., and Johnson, C. 2009. *Animals make us human : creating the best life for animals*. Houghton Mifflin Harcourt, Boston.
- Haft, D.H., Selengut, J., Mongodin, E.F., and Nelson, K.E. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol* **1**: e60. doi: 10.1371/journal.pcbi.0010060.
- Hayashi, M., Maehara, K., Harada, A., Semba, Y., Kudo, K., Takahashi, H., et al. 2016. Chd5 Regulates MuERV-L/MERVL Expression in Mouse Embryonic Stem Cells Via H3K27me3

- Modification and Histone H3.1/H3.2. *J Cell Biochem* **117**: 780-792. doi: 10.1002/jcb.25368.
- Hsu, P.D., Lander, E.S., and Zhang, F. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**: 1262-1278. doi: 10.1016/j.cell.2014.05.010.
- Huisman, C., van der Wijst, M.G., Falahi, F., Overkamp, J., Karsten, G., Terpstra, M.M., et al. 2015. Prolonged re-expression of the hypermethylated gene EPB41L3 using artificial transcription factors and epigenetic drugs. *Epigenetics* **10**: 384-396. doi: 10.1080/15592294.2015.1034415.
- 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.
- Jansen, R., Embden, J.D., Gaastra, W., and Schouls, L.M. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**: 1565-1575.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816-821. doi: 10.1126/science.1225829.
- Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. 2014. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* **24**: 1012-1019. doi: 10.1101/gr.171322.113.
- Kim, Y.G., Cha, J., and Chandrasegaran, S. 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* **93**: 1156-1160.
- Kunin, V., Sorek, R., and Hugenholtz, P. 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome biology* **8**: R61. doi: 10.1186/gb-2007-8-4-r61.
- Lander, E.S. 2015. Brave New Genome. *N Engl J Med* **373**: 5-8. doi: 10.1056/NEJMp1506446.
- Laufer, B.I., and Singh, S.M. 2015. Strategies for precision modulation of gene expression by epigenome editing: an overview. *Epigenetics Chromatin* **8**: 34. doi: 10.1186/s13072-015-0023-7.
- Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., et al. 2014. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome biology* **15**: 554. doi: 10.1186/s13059-014-0554-4.
- Liang, P., Xu, Y., Zhang, X., Ding, C., Huang, R., Zhang, Z., et al. 2015a. CRISPR/Cas9-mediated gene editing in human triploid zygotes. *Protein Cell* **6**: 363-372. doi: 10.1007/s13238-015-0153-5.
- Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., et al. 2015b. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J Biotechnol* **208**: 44-53. doi: 10.1016/j.jbiotec.2015.04.024.
- Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., et al. 2016. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**: 400-403. doi: 10.1126/science.aad5725.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., et al. 2011. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**: 467-477. doi: 10.1038/nrmicro2577.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., et al. 2015. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**: 722-736. doi: 10.1038/nrmicro3569.
- Marraffini, L.A., and Sontheimer, E.J. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* **11**: 181-190. doi: 10.1038/nrg2749.

- Millman, S.T., Duncan, I.J., and Widowski, T.M. 2000. Male broiler breeder fowl display high levels of aggression toward females. *Poult Sci* **79**: 1233-1241.
- Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**: 174-182. doi: 10.1007/s00239-004-0046-3.
- Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzone, E.O., et al. 2015. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* **347**: 1017-1021. doi: 10.1126/science.1262088.
- Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., et al. 2016. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**: 403-407. doi: 10.1126/science.aad5143.
- Pourcel, C., Salvignol, G., and Vergnaud, G. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**: 653-663. doi: 10.1099/mic.0.27437-0.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., et al. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**: 1173-1183. doi: 10.1016/j.cell.2013.02.022.
- Ramalingam, S., Annaluru, N., and Chandrasegaran, S. 2013. A CRISPR way to engineer the human genome. *Genome biology* **14**: 107. doi: 10.1186/gb-2013-14-2-107.
- Ramanan, V., Shlomai, A., Cox, D.B., Schwartz, R.E., Michailidis, E., Bhatta, A., et al. 2015. CRISPR/Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus. *Sci Rep* **5**: 10833. doi: 10.1038/srep10833.
- Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., et al. 2013a. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**: 1380-1389. doi: 10.1016/j.cell.2013.08.021.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. 2013b. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**: 2281-2308. doi: 10.1038/nprot.2013.143.
- Rath, D., Amlinger, L., Rath, A., and Lundgren, M. 2015. The CRISPR-Cas immune system: biology, mechanisms and applications. *Biochimie* **117**: 119-128. doi: 10.1016/j.biochi.2015.03.025.
- Sung, Y.H., Kim, J.M., Kim, H.T., Lee, J., Jeon, J., Jin, Y., et al. 2014. Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res* **24**: 125-131. doi: 10.1101/gr.163394.113.
- Tabebordbar, M., Zhu, K., Cheng, J.K., Chew, W.L., Widrick, J.J., Yan, W.X., et al. 2016. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**: 407-411. doi: 10.1126/science.aad5177.
- van Erp, P.B., Bloomer, G., Wilkinson, R., and Wiedenheft, B. 2015. The history and market impact of CRISPR RNA-guided nucleases. *Curr Opin Virol* **12**: 85-90. doi: 10.1016/j.coviro.2015.03.011.
- Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., et al. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**: 910-918. doi: 10.1016/j.cell.2013.04.025.
- Webber, B.L., Raghu, S., and Edwards, O.R. 2015. Opinion: Is CRISPR-based gene drive a biocontrol silver bullet or global conservation threat? *Proc Natl Acad Sci U S A* **112**: 10565-10567. doi: 10.1073/pnas.1514258112.
- Wu, Y., Zhou, H., Fan, X., Zhang, Y., Zhang, M., Wang, Y., et al. 2015. Correction of a genetic disease by CRISPR-Cas9-mediated gene editing in mouse spermatogonial stem cells. *Cell Res* **25**: 67-79. doi: 10.1038/cr.2014.160.

- Yang, W., Tu, Z., Sun, Q., and Li, X.J. 2016. CRISPR/Cas9: Implications for Modeling and Therapy of Neurodegenerative Diseases. *Front Mol Neurosci* **9**: 30. doi: 10.3389/fnmol.2016.00030.
- Yuan, M., Gao, X., Chard, L.S., Ali, Z., Ahmed, J., Li, Y., et al. 2015. A marker-free system for highly efficient construction of vaccinia virus vectors using CRISPR Cas9. *Mol Ther Methods Clin Dev* **2**: 15035. doi: 10.1038/mtm.2015.35.
- Zhai, X., Ng, V., and Lie, R. 2016. No ethical divide between China and the West in human embryo research. *Dev World Bioeth* **16**: 116-120. doi: 10.1111/dewb.12108.
- Zhang, F. 2015. CRISPR-Cas9: Prospects and Challenges. *Hum Gene Ther* **26**: 409-410. doi: 10.1089/hum.2015.29002.fzh.

Draft

Figure Legends

Figure 1. The phases of CRISPR Cas systems. CRISPR Cas systems work in three phases; acquisition, expression and inference. In type I the Cascade system is responsible for cleaving protospacers, where type II uses Cas9, and type III uses Cascade like systems, Csm or Cmr (Makarova et al. 2011). Permission to re-use Figure 1 in (Makarova et al. 2011) given by Nature Publishing group.

Figure 2. Overview of CRISPR Cas9 system protocol for genome editing (Ran et al. 2013b). Permission to re-use Figure 2 in (Ran et al. 2013b) given by Nature Publishing group.