

A Review of CRISPR-Based Genome Editing: Survival, Evolution and Challenges

Hafiz Ishfaq Ahmad¹, Muhammad Jamil Ahmad¹, Akhtar Rasool Asif², Muhammad Adnan³, Muhammad Kashif Iqbal⁴, Khalid Mehmood^{4,5}, Sayyed Aun Muhammad², Ali Akbar Bhuiyan^{1,6}, Abdelmoteleb Elokil^{1,7}, Xiaoyong Du¹, Changzhi Zhao¹, Xiangdong Liu^{1*} and Shengsong Xie^{1,4*}

¹Key Laboratory of Agricultural Animal Genetics, Breeding, and Reproduction of Ministry of Education and Key Laboratory of Swine Genetics and Breeding of Ministry of Agriculture, College of Animal Science and Technology, Huazhong Agricultural University, Wuhan 430070, P. R. China;

²University of Veterinary and Animal Sciences, Lahore, sub campus, Jhang 35200, Pakistan;

³College of Plant Science, Huazhong Agricultural University, Wuhan 430070, P. R. China;

⁴College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, P. R. China;

⁵University College of Veterinary and Animal Sciences, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan;

⁶Senior Scientific Officer, Bangladesh Livestock Research Institute, Savar, Dhaka 1341, Bangladesh;

⁷Animal Production Departments, Faculty of Agriculture, Benha University, Moshtohor 13736, Egypt.

*Correspondence: ssxie@mail.hzau.edu.cn;
liuxiangdong@mail.hzau.edu.cn

DOI: <https://doi.org/10.21775/cimb.028.047>

Abstract

Precise nucleic acid editing technologies have facilitated the research of cellular function and the development of novel therapeutics, especially the current programmable nucleases-based editing tools, such as the prokaryotic clustered regularly interspaced short palindromic repeats (CRISPR)-associated nucleases (Cas). As CRISPR-based therapies are advancing toward human clinical trials, it is important to understand how natural

genetic variation in the human population may affect the results of these trials and even patient safety. The development of "base-editing" technique allows the direct, stable transformation of target DNA base into an alternative in a programmable way, without DNA double strand cleavage or a donor template. Genome-editing techniques hold promises for the treatment of genetic disease at the DNA level by blocking the sequences associated with disease from producing disease-causing proteins. Currently, scientists can select the gene they want to modify, use the Cas9 as a "molecular cutter" to cut it out, and transform it into a more desirable version. In this review, we focus on the recent advances of CRISPR/Cas system by outlining the evolutionary and biotechnological implications of current strategies for improving the specificity and accuracy of these genome-editing technologies.

1. Introduction

For many years, strategies for efficient genome editing and targeted alterations were restricted to specific creatures (e.g., homologous recombination in yeast or mice) with the prerequisite of selectable markers for drug modification methods (Sander and Joung 2014). The introduction of targeted genome editing into living organisms has become a powerful tool for biology and medicine (Dow et al., 2015). Frame shift mutation knockout enables precise genome editing such as insertion, deletion, specific sequence alteration, induction of amino acid substitution for disease modeling, incorporation of traits of interest into crops and animals, and modification of genetic defects for therapeutic purpose in most of the organisms and cell types (Sander and Joung 2014). Thomas and coworkers have laid the foundation of genetic research by inducing targeted genome mutation to produce knockout mice (Thomas and Capecchi 1987). The long-cherished idea of genome modification has resulted in advanced, efficient and precise genome manipulation techniques (Dow and Lowe 2012). Both zinc-finger nucleases (ZFNs) and transcription activator-like effectors nucleases (TALENs) involve

the development of chimeric proteins against target DNA sequence. However, the new era of genetic engineering came with the emergence of CRISPR, which made possible the generation of double strand breaks (DSBs) of a target DNA locus (Boch et al., 2009; Cermak et al., 2011; Moscou and Bogdanove 2009; Ran et al., 2013; Urnov et al., 2010). While these techniques involve various mechanisms, all of them perform the function of genome editing by inducing DSBs (Bosch et al., 2002). The key characteristics to targeted genetic engineering technologies are the targeted DNA damage and non-homologous end joining (NHEJ) DNA repair or homology directed repair (HDR) (Dow et al., 2015). HDR provides high fidelity repair by allowing the incorporation of a foreign donor DNA as a repairing template. All the vector-based genome editing technologies utilize HDR approach, and rely on less DNA destruction frequency at the target point (Kim and Kim 2014). However, ZFNs, TALENs and CRISPRs are prone to potential targeted genome disruption by NHEJ or efficient HDR-mediated DNA DSB breaks at target sites. Several researchers have identified DNA bias repair inhibitors towards either HDR or NHEJ, providing control strategies over genome revision (Chu et al., 2015; Maruyama et al., 2015). Despite a good review of TALEN and ZFN-based genomic engineering approaches elsewhere (Gaj et al., 2013), neither of them has attained the purpose of quick and accurate genome editing that was initially hyped due to limitations in generation of reagents (ZFN or TALEN) (Hockemeyer et al., 2011; Miller et al., 2011; Wood et al., 2011). Furthermore, the ZFN methodology itself is also problematic and the TALEN system is provided as an alternative open source (Cermak et al., 2011; Sanjana et al., 2012). Since a short demonstration in 2013, the CRISPR system has taken the scientific world as a potential genome editing tool (Doudna and Charpentier 2014; Mali et al., 2013a; Mali et al., 2013b) to cleave DNA sequence (Bolotin et al., 2005; Jansen et al., 2002). Soon after that, the genome editing revolution brought about new technology innovations, enabling silent mutation to induce and replace the genomic elements (Cong et al., 2013) and the CRISPR power was quickly revealed by genome modification in mammalian cells (Jinek et al., 2013). When compared to TALEN and ZFN techniques, CRISPR (Cas) technology is more efficient and simple for designing RNA-guided molecules (Friedland et al., 2013). Therefore, CRISPR has been rapidly adopted as the genome-editing system in a short time and applied to all kinds of living organisms (Hruscha et al., 2013). In this review, we describe

the CRISPR/Cas system in terms of evolution, diversity, recent research made possible by CRISPR/Cas technologies as well as their advantages and limitations. We also discuss applications and current strategies for improving the specificity and accuracy of these technologies. Finally, we propose the challenges in this emerging genomic engineering platform that need to be addressed in the future.

2. Programmable Genome Editing - The goal of a decade

Ever since the discovery of the DNA double strand, researchers have been speculating about the possible site-specific changes to cell or organism genome as the earliest approaches worked on the principle of site-specific DNA sequence recognition. The study in yeast and bacteria clarified endogenous machinery within the cells to repair DNA double strand breaks naturally (Doudna and Charpentier 2014), thus precise site-specific DNA break methods have been introduced as a valuable strategy for programmable genome engineering. Earlier approaches used oligonucleotides to recognize DNA base to target the DNA cleavage (Varshavsky et al., 2006). Cross linking between oligonucleotides and reagents such as psoralen and bleomycin was found appropriate to site-specific genome editing in mammalian and yeast cells (Faruqi et al., 1996; Sandor and Bredberg 1995; Strobel and Dervan 1990, 1991). The bleomycin was used as a chemical method for DNA sequence recognition that enables the binding of targeted chromosomal loci and can be modified by coupling chemical recognition reagent with cleavage reagent (Cho et al., 1995; Faruqi et al., 1998; Gottesfeld et al., 1997). Furthermore, self-splicing introns-based nucleic acid base pairing approach was used to change DNA or RNA sequences (Yang et al., 1996; Zimmerly et al., 1995). These strategies, despite demonstrating the site-specific genome modifications, could not lead to robust methods. Similarly, the combination of zinc finger mediated DNA binding domain (Pavletich and Pabo 1991) with the nuclease of the restriction enzyme *FokI* domain independently led to the development of modular DNA recognition protein, which could function as the site-specific nuclease (Kim et al., 1996). Zinc finger nucleases (ZFNs) were used effectively in mammalian cells and drosophila for induction of genomic sequence changes (Bibikova et al., 2003; Bibikova et al., 2002). ZFNs could not be widely used due to its difficulty to formulate and validate site specific DNA recognition proteins, thus the field was primed for reporting bacterial and plant

TALENs, enabling rapid creation of *FokI* coupled version similar to ZFNs for targeted genome editing (Boch et al., 2009; Christian et al., 2010; Moscou and Bogdanove 2009). These chimeric nucleases enable precise and efficient genome modification by stimulating DNA repair mechanisms within the cell with NHEJ and HDR through targeted DSB induction (Wyman and Kanaar 2006). The protein byproducts of ZFNs and TALENs act as DNA binding spheres to adapt this technology. Such simple and flexible combination has boosted ZFNs and TALENs to the forefront of programmable genome engineering (Gaj et al., 2013). The most common motif for DNA ligation, Cys₂-His₂ zinc finger was found in eukaryotes and corresponded to the frequently programmed protein domain in the human genome (Beerli and Barbas 2002). Conserved $\beta\beta\alpha$ configuration, consisting of around 30 amino acids in zinc finger protein, enables specific DNA recognition. The structure-based advancement in highly conserved linkers facilitated the development of synthetic zinc finger proteins that recognize 9-18 bp length DNA sequences (Liu et al., 1997), because these domains can recognize all of nucleotide triplets known to date and zinc finger components selected previously could be connected in tandem to DNA sequences containing the sequences of these DNA triplets (Bhakta et al., 2013; Gonzalez et al., 2010; Kim et al., 2011). Zinc finger protein domains have been the single tactic to construct DNA binding enzymes and site-specific proteins (Gaj et al., 2013). The discovery of the TALEN protein has provided the alternative platform for programmable genome engineering (Moscou and Bogdanove 2009). TALENs hold DNA binding domains consisting of 33-35 amino acids repeats, which are naturally present in plant pathogenic bacterial *Genus Xanthomonas* (Deng et al., 2012). There was no linkage reengineering between TALEN repeats with single site targeting genome ability similar to that of zinc finger proteins. In the past decade, for the revolutionary technology of targeted genetic modification founded by zinc finger protein, several effector domains have been used to fuse with the recurrences of TALENs including transcription like effectors (Zhang et al., 2011), nucleases (Miller et al., 2011; Mussolino et al., 2011) and site-specific recombinases (Mercer et al., 2012).

3. Evolution of CRISPR/Cas systems - Roots to Riches

Since the beginning of phage study in laboratory (d'Herelle et al., 1917; Twort et al., 1915), researchers have identified and categorized the

bacteria resistant to phage infection (Burnet et al., 1929; Luria and Delbrück 1943), resulting in the discovery of different defense mechanisms against viruses. A common defense mechanism against phage is inhibition of phage adsorption or genome injection. Bacteria are capable of switching on or off the phage receptor (Zaleski et al., 2005) manifestation or secretion of polysaccharides that limit receptor access (Hanlon et al., 2001). The most common method for analysis of phage infection is based on sensor protein to activate membrane depolarization (Parma et al., 1992) or bacterial cell translation inhibition that leads to cell death (Bingham et al., 2000). The defense system of prokaryotes named restriction modification uses nucleases to identify and cleave the short DNA motif in the invading genome to ensure the protection (Bickle and Krüger 1993). Blockage of phage replication and attack of foreign invading elements are two other defense systems in the Bacteriophage exclusion system. The host population can restore the missing defense mechanism to adjust phage invasion strategies by selection of infrequent mutations. By contrast, the introduction of targeted site mutations using spacer sequences in CRISPR loci provides a quick and effective defense response against phages (Childs et al., 2014; Paez-Espino et al., 2015; Weinberger et al., 2012). Distinctive defense mechanisms are encoded by CRISPR loci and associated (Cas) genes, leading to robust adaptation to the speedily developing dsDNA viruses or other foreign DNA viruses of archaeobacteria. CRISPR loci consist of short (30-40 bp) palindromic, repetitive sequences interspaced by short spacer sequences from viral origin plasmid, and they specify immunity against phages or plasmids containing complementary sequences, which is central to CRISPR defense (Barrangou et al., 2007; Marraffini and Sontheimer 2008). The anti-viral or anti-plasmid immunity involved in CRISPR/Cas loci was initially evaluated in-silico through its first description in 1987 after sequencing *Escherichia coli* gene (Ishino et al., 1987). Distinct arrays of short repeats interspaced with unique spacer sequences have been recognized in bacterial and archaeal genome and these short arrays were anticipated to have an important function (Mojica et al., 2000). Independently, Cas proteins adjacent to CRISPR sequences encoded by putative operons were studied in detail using comprehensive methods and found to contain domains that are featured by several nucleases, a polymerase, helicase and different RNA binding proteins (Jansen et al., 2002). These proteins were initially supposed to constitute

a novel DNA repairing system (Makarova et al., 2002), but observations established a hypothesis that CRISPR/Cas system involves the defense against invading selfish elements (Mojica et al., 2005) because of the unique CRISPR spacer sequences which are almost identical to those of viral or plasmid gene fragments (Bolotin et al., 2005; Pourcel et al., 2005). Based on these findings and other computational analyses of Cas proteins (Haft et al., 2005), an analogy model was proposed (Makarova et al., 2006) between the CRISPR/Cas system of bacteria and archaea and the eukaryotic RNA interference (RNAi) mechanism (Carthew and Sontheimer 2009). The ancestral CRISPR/Cas may belong to class I system due to the widespread existence of this system in bacteria and archaea, together with antique RRM domain in class I effector protein proliferation. The class II system variants have most likely evolved through various independent substitutions of class I effectors with nuclease genes that were derived from unique mobile genetic elements (MGEs). The type V effector variants (Cpf1) may have evolved from different TnpB transposase gene families that are widespread in transposons (Shmakov et al., 2015), while the type II effector may have evolved from IscB, which is a two nuclease domain protein belonging to newly recognized transposons family (Kapitonov et al., 2016). Both Cas1 and Cas2 genes of CRISPR/Cas systems are responsible for spacer procurement, and both of them are fundamentally uniform. CRISPR/Cas component exhibits an outlying versatility, due to its involvement in development of CRISPR RNA (crRNA) as well as marked recognition and cleavage (Makarova et al., 2015).

4. CRISPR adaptation as a defense system

The spacers of CRISPR/Cas systems represent a sequential archive of former intruder encounters. The detained spacer sequence is incorporated into the CRISPR locus after interaction to MGEs, and provide defense against MGEs, particularly viruses and plasmids, by sequences specifically targeting invading DNA or RNA (Barrangou et al., 2007; Terns and Terns 2011). The invader sequence analysis disclosed a short motif next to the target sequence, named protospacer adjacent motif (PAM) (Deveau et al., 2008), and these spacers are the main units of adaptive immunity, as they can hold memory of invaders with specific MGEs acquired from previous infection, which enables the recognition and counteraction of intruders upon subsequent infection (Barrangou et al., 2007). These PAM sequences allow the self or non-self-recognition by

the host in various ways such as the presence of PAM sequences in foreign targets, which is obligatory for the efficiency of target recognition, and the absence of PAM sequence in the host system prevents self-targeting (Mojica et al., 2009). In type I and type II systems, in addition to participation in interference, PAM plays a major part in spacer selection throughout adaptation phase, involving the procurement of functional spacers (Datsenko et al., 2012). The PAM is a short sequence that cannot prevent the integration of spacers itself from host DNA due to less information of the motif. The PAM emerges from an evolutionary adjustment between competent integration of spacers from non-self DNA and preclusion of an autoimmune response (Mohanraju et al., 2016). Cas1 and Cas2 proteins play a vital role in spacer procurement in all systems (Yosef et al., 2012). However, Cas1 and Cas2 genes are absent in several CRISPR/Cas loci, particularly type III, type IV and type VI (Makarova et al., 2015). Mutations in Cas1 active site eliminate spacer incorporation in *E. coli*, while the Cas2 nuclease activity is expendable. In *E. coli*, a Cas2 dimer and two Cas1 dimers form a composite, which fuse and processes the PAM comprising DNA fragments (Nuñez et al., 2014), followed by incorporation of the newly produced spacers into CRISPR array by recombination tool similar to that of retroviral transposases and integrases (Nuñez et al., 2015).

Adaptation (spacer acquisition) in type I system continues along two pathways: a naïve procurement that proceeds during an early infection and a primed integration, which is complementary to the foreign DNA and occurs when CRISPR contains previously integrated spacer (Fineran and Charpentier 2012). Conferring to the suggested model, naïve spacer acquisition involves five phases: (i) fragmentation of intrusive DNA by RecBCD or by restriction endonucleases (Levy et al., 2015), (ii) selection of protospacer for PAMs by one of the subunits of Cas1-Cas2 complex (Arslan et al., 2014), (iii) evaluation of selected protospacers using 3-hydroxyl groups by Cas1 nucleases, (iv) nicking of both leader proximal strands of CRISPR array at 5' end by direct nucleophilic attack, followed by covalent linkage of each of the newly selected spacer strands to the single stranded duplication ends, and (v) construction and ligation of the second strand of repeat flanked by a non-Cas repair method (Figure 1) (Amitai and Sorek 2016; Silas et al., 2016). Primed acquisition constitutes a positive feedback mechanism, which facilitates new spacer acquisition from previously encountered genetic components (Fineran et al., 2014). Priming

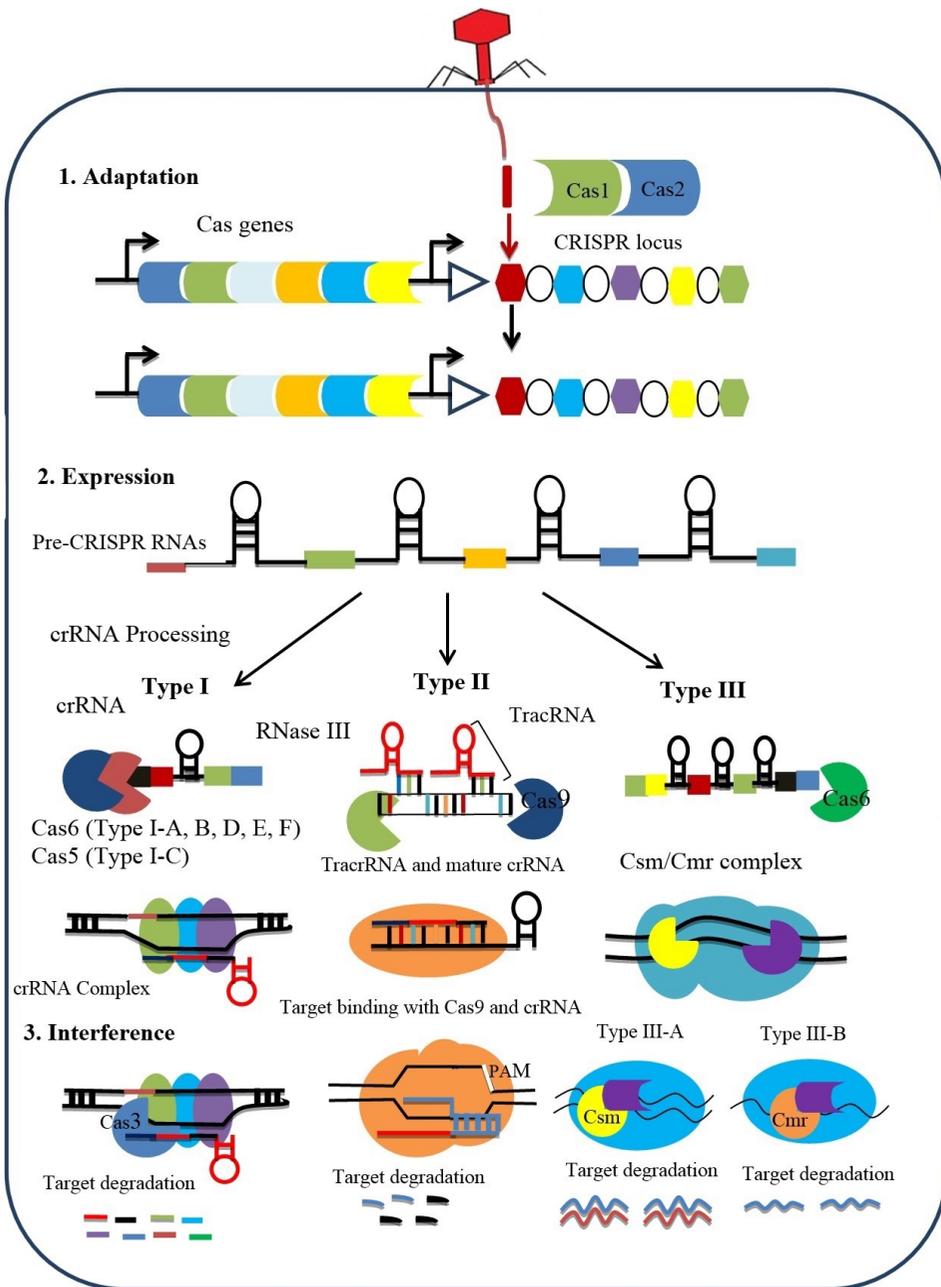


Figure 1. Mechanism of CRISPR immunity in different types of CRISPR/Cas classifications. In type I classification, the crRNA precursor (pre-crRNA) inside the cascade complex is handled by cas6 endoribonuclease after spacer acquisition in the presence of Cas1 and Cas2 proteins. Mature crRNA (mat-crRNA) is released at the base of stem-loop designed by repeat RNA cleavage. Next, the cascade complex containing the mat-crRNA assembles the cas3 nuclease for cutting the DNA strand complementary to the protospacer, directly downstream of the interaction region through crRNA spacer. This leads to the collapse of a foreign DNA. In type II classification, after spacer acquisition, tracrRNA is used for the biosynthesis of crRNA. The small crRNA is released by coupling between pre-crRNA and tracrRNA sequences in the development of a double strand substrate, and cleavage by the host-encoded RNase III in the presence of cas9. After the release of the intermediate crRNA (int-crRNA), there is instant trimming of int-crRNA at 5' end that produces mat-crRNA. For target cleavage, tracrRNA, crRNA, and cas9 domains (Ruvc and HNH) are required. The DNA strand is cleaved through the HNH domain that is complementary to crRNA guide, whereas non-complementary strand is cleaved through Ruvc domain. In type III classification, mat-crRNA is released by the transformation of pre-crRNA to int-crRNA in the presence of cas6, followed by integration of int-crRNAs into Cmr/cas10 or Csm/cas10 complex, where additional maturation events occur by trimming at 3' end sequences. Genetic evidence showed that DNA was cleaved by sub-type III-A, whereas subtype III-B was shown to help the cleavage of the RNA molecules according to the biochemical data.

seems to be the major pathway of type I CRISPR adaptation and is strongly dependent on spacer sequences (Li et al., 2014; Xue et al., 2015).

5. CRISPR/Cas proteins and their functional modules

In addition to an instant evolution in Cas genes, the prominent variability of Cas loci in the genome architecture poses a major contest for reliable footnoting of Cas proteins and for CRISPR/Cas classification (Vestergaard et al., 2014; Wang et al., 2011). Therefore, a stable classification array is essential for accurate and durable categorization of CRISPR/Cas loci in genomes. Because of complexity in gene composition, all the sorting principles were squeezed, and phylogenetic methodology was developed on the basis of genomic structural analysis. In classification hierarchy, six types of CRISPR/Cas systems are rapidly distinguished by virtue of unique proteins signature, which are further grouped into two classes. The up-to-date classification for CRISPR/Cas system takes into account of the stock of the locus architecture and the sequence similarity between Cas proteins, including two classes that are currently subdivided into six types and nineteen subtypes (Makarova et al., 2015; Shmakov et al., 2015). The two classes of CRISPR/Cas proteins vary primarily in the organization of effector modules. Class I system includes types I, III, and IV which are present in archaea and bacteria and comprise effector complexes containing four to seven Cas protein subunits in an irregular organization (e. g., complexes of type I system involved in antiviral defense and Cmr/Csm complexes of type III system. Subunits Cas5, Cas6 and Cas7 of Class I have RNA mandatory RRM domains (RNA recognition motifs) (Figure 1) (Makarova et al., 2011; Makarova et al., 2013). Despite low resemblance of effector complex subunits between type I and type III, their identical architecture demonstrates a common origin (Rouillon et al., 2013; Spilman et al., 2013).

Class II CRISPR/Cas system includes types II, V, and VI, which are effectively restricted to bacteria, and Cas9, the RNA dependent endonuclease, is the most efficiently characterized effector. Cas9 consists of two distinct nuclease domains, RuvC and HNH, which are responsible for the cleavage of displaced and targeted strands in the crRNA-target DNA complex respectively (Jinek et al., 2012). The type II system also encodes trans-acting crRNA that evolved from particular CRISPR repeats and it is necessary for pre-CRISPR RNA processing and

target recognition in type II system (Figure 1) (Chylinski et al., 2014; Deltcheva et al., 2011). The phenotype of type V effector and its subtype (Cpf1) contains only one (RuvC like) domain that is identified by sequence analysis, but the recent analysis revealed the second nuclease domain in the structure of Cpf1 is fused with the target DNA and crRNA which is distinct from HNH or other nucleases (Zetsche et al., 2015). Screening of metagenomes and microbial genomes has identified three novel CRISPR/Cas variants, including subtypes V-B and V-C, which resemble Cpf1 in that their effector proteins contain a single RuvC-like nuclease domain. Target DNA cleavage by the type V-B effector, denoted as C2c1, and has been experimentally validated (Zetsche et al., 2015). Type VI is distinct from the other types in that its effector proteins contain two preserved HEPN areas that own ribonuclease activity. Current comparative studies of CRISPR/Cas system variants have discovered a robust integrated evolution with several arrangements of effector modules, and adaptation modules as well as a focal influence of MGEs on origin and divergence of CRISPR/Cas systems (Shmakov et al., 2015).

The Cas proteins are part of the four discrete functional modules: spacer acquisition (adaptation), CRISPR RNA processing and target binding (expression), target cleavage (interference), and regulatory and other CRISPR associated functions (ancillary) (Makarova et al., 2015). In modern era, the central Cas proteins (Cas1-Cas10) have evidently unique structures and functions, enabling them to be categorized into various modules (Makarova et al., 2015). The constant module spacer acquisition across CRISPR system contain Cas1 and Cas2 proteins influences the Cas4 endonuclease enzyme (Hooton and Connerton 2015), and Cas9 in type II CRISPR system (Wei et al., 2015). The integrase Cas1 is able to adopt a distinct α -helical fold and induce cleavage at particular sites within the CRISPR system by insertion of unique spacers (Nuñez et al., 2014; Nuñez et al., 2015). Cas2, a homolog to mRNA interferase, forms a composite with Cas1 in type I systems and is essential for adaptation. The expression and intrusion modules are characterized by using multi-subunit crRNA-effector complex (Mulepati et al., 2014; Taylor et al., 2015), or in type II system, by a large single protein, Cas9 (Jinek et al., 2014). In the expression phase, a pre-crRNA is attached to a multi-subunit crRNA-effector complex, or to Cas9, and changed into a complete crRNA

catalyzed by RNA endonucleases (Wang et al., 2011; Zhang et al., 2013).

With reference to interference module of type I and type II systems, the crRNA effector complex or Cas9 pools nuclease activity with RNA binding domains through base pair formation with crRNA spacer region (Makarova et al., 2011). Target cleavage is catalyzed by Cas3 or its domains (HD family nucleases) in type I system by joint action of Cas7 and Cas10 proteins in type III system (Osawa et al., 2015; Taylor et al., 2015). The HD nuclease domains are either fused to superfamily helicase 2 or are encoded by a single gene, Cas3, in type I system, where, just like in type III system, HD nuclease domains are fused to Cas10 and slice ssDNA during interference phase (Jung et al., 2015). In type II system, the RuvC-like nuclease domains and the HNH nuclease domains of Cas9 each slice one of the target DNA strands (Jinek et al., 2013). Remarkably, a large multi-domain Cas9 protein is prerequisite for all the three (adaptation, expression and interference) functional modules of CRISPR-based immunity in type II system (Makarova et al., 2015).

The auxiliary unit is the arrangement of several proteins including spheres except Cas4, which are less than the CRISPR/Cas system proteins. The auxiliary unit not only has an assumed role in adaptation but is also supposed to promote the contribution of Cas4 to CRISPR/Cas mediated encoded cell death (Makarova et al., 2012). Further important parts in the auxiliary unit are an CRISPR-Associated-Rossmann-Fold (CARF) domain with a distinct group of proteins (Makarova et al., 2014) that have been postulated to control CRISPR/Cas activity, and the deactivated P-loop ATPase-Csn2 that makes a homo-tetrameric loop to accommodate linear double-stranded DNA (dsDNA) in the central region (Lee et al., 2012). A novel hyper-accurate Cas9 variant (HypaCas9) was found, which reveal high genome wide specificity without on target activity adjustment in human cells. Sequence investigation exposed that HypaCas9 displays enhanced genome wide specificity corresponding with wild-type SpCas9, and exhibited equal or improved genome wide specificity comparative to eSpCas9 and SpCas9-HF1 for all single guide RNAs (sgRNAs). These characteristics validate the improved mismatch bias of HypaCas9 and establish that its specificity enhancements may range beyond the PAM middle and proximal section of the spacer sequence. These outcomes defined a universal approach for refining the Cas9 specificity by

modification of natural configurational threshold, and provide an opportunity for rational design of hyper-accurate Cas9 variants, which do not compromise efficiency (Chen et al., 2017).

6. Prospective applications of CRISPR/Cas9 in genome engineering

i. Cas9 targeting revolution

In the most recent era, the prokaryotic CRISPR/Cas9 system-based genome editing approach has transformed genome editing through its unique efficacy and remarkable simplicity in use (Figure 2A). CRISPR/Cas9 editing has replaced the application of TALENs and ZFNs in cell types and a multitude of model organisms (Barrangou et al., 2007). The Cryo-electron microscopy (cryo-EM) technology has proved that Cas9/sgRNA-target DNA complex provides an efficient molecular mechanism and the basis for engineering an upgraded CRISPR/Cas9 (Slaymaker et al., 2016). Studies have revealed the detailed molecular interaction between sgRNA and CRISPR/Cas9 as well as target DNA and CRISPR/Cas9 (Jinek et al., 2012). A unique Cas9 bilobed architecture consists of a nuclease domain and a sgRNA binding target recognition lobe and is supposed to carry the PAM interaction domain. A groove formed between the two lobes has sgRNA target DNA heteroduplex. The configuration of these interactions will facilitate structure-based rational design of optimal guides by defining the elements that drive PAM binding fidelity, mismatch tolerance, and specificity for different Cas9 protein systems, along with resemblances and variations among different CRISPR/Cas systems (Barrangou et al., 2007). Major features underlying the CRISPR/Cas9 genome editing revolution include the compactness, simplicity and targeting flexibility afforded by this system (Sander and Joung 2014). The implications for Cas9 enhanced targeting specificity and binding capability are exciting and have laid a basis for further optimization of the Cas9 system and the improvement of next-generation CRISPR tools. Forthcoming work will determine the potential of different engineered or natural Cas9 proteins, sgRNAs, and their relevant PAM sequences for enhanced sequence recognition and target DNA binding efficacy.

ii. Editing out human diseases

Tyrosinaemia, a human metabolic disease, was first corrected in mice by use of CRISPR to fix the disease-initiating mutations in adult animals, which was an important step towards gene therapy in humans using CRISPR technology (Yin et al.,

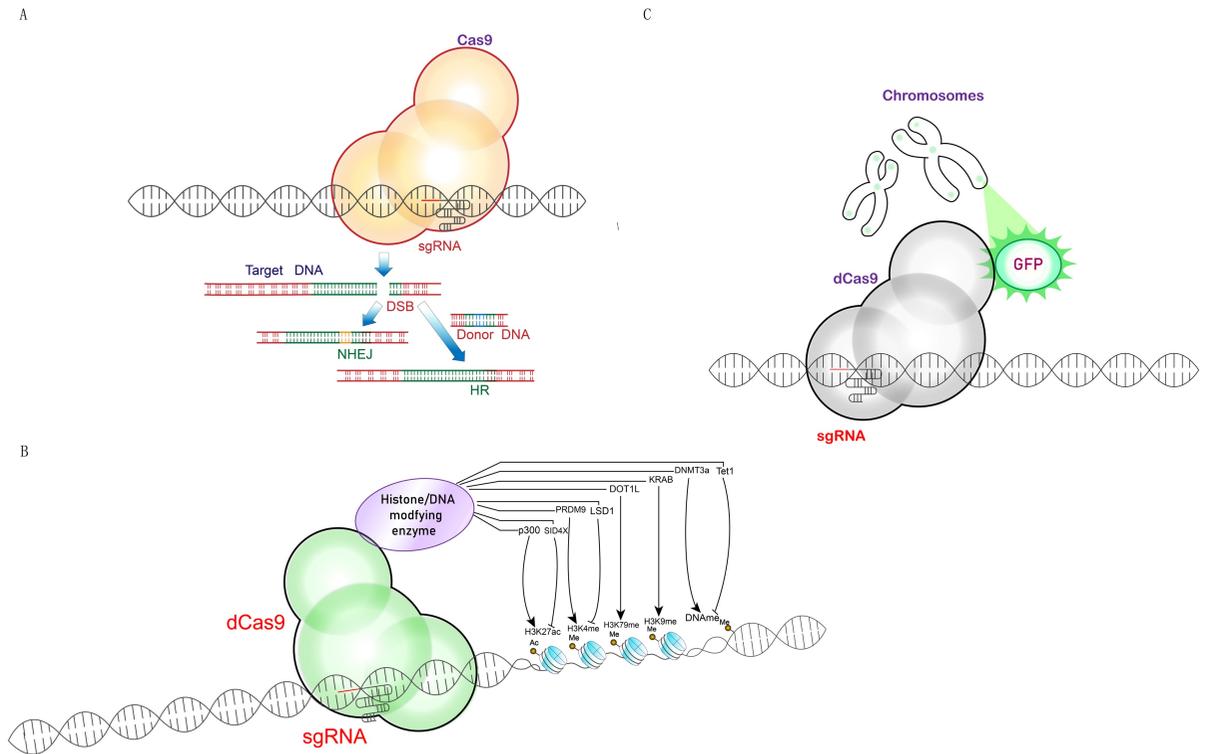


Figure 2. Applications of CRISPR/Cas9 technology. **A. Genome editing:** CRISPR/Cas9 is used for making double strand breaks (DSB) at particular sites. These introduced DSBs are mended by either non-homologous end-joining (NHEJ) or homology-directed repair (HDR) (the cartoon on the left). Efficient reorganization over a long distance in the genome can be endorsed between the pairs of targeted loci (the cartoon on the right). **B. Epigenetic modulation:** DNA methylation and histone modification; dCas9-LSD1 carries histone methylation for gene expression repression precisely and p300-dCas9 carries histone acetylation that brings gene expression activation precisely. Epigenetic effectors merged with dCas9 could change epigenetic states of the target loci, such as DNA methylation or histone modifications. **C. Genomic imaging:** dCas9 was merged with a fluorescent protein to envisage specific genome loci in live cells. DSB stands for DNA double-strand breaks, NHEJ stands for non-homologous end joining, HR stands for homologous recombination, dCas9 stands for catalytically inactive CRISPR-associated nuclease 9, GFP stands for green fluorescent protein, sgRNA stands for single guide RNA.

2014). Hereditary tyrosinemia type I disorder is caused by lack of fumarylacetoacetate hydrolase, leading to cell death of hepatocytes due to accumulation of cytotoxic metabolite. The specific sgRNA and the vectors encoding Cas9 together with a DNA oligo for HDR were delivered directly into the mouse liver through hydrodynamic tail vein injection, which corrected the fumarylacetoacetate hydrolase allele mutations and rescued the weight loss of mice by reducing hepatocellular toxicity (Yin et al., 2014).

Despite the development in antiviral remedies for treating patients with Chronic HBV, the progress of HBV infection into chronic hepatitis resulted in hepatocellular carcinoma or liver cirrhosis in many patients. Notably, CRISPR/Cas9 system co-injection

targets the sequence directed to cleavage of HBV expression vector, resulting in the reduction of serum HBV surface antigen levels. In experiments similar to those of a human cell line, scientists eradicated covalently closed circular DNA (cccDNA) of duck HBV, achieving a complete elimination of HBV by genome excision (Lin et al., 2014). The genetic blood disorder beta thalassemia was cured by using CRISPR/Cas9-based gene therapy in induced pluripotent stem cells (iPSCs). The iPSCs were established from fibroblasts of a beta thalassemia patient and were transfected through CRISPR/Cas9-based vectors to target the disease-causing allele (Xie et al., 2014).

Furthermore, the potential applications of CRISPR/Cas9 are not only restricted to therapeutic manipulations of the cancer genome. About 16.1% of cancer occurrence was estimated to be caused by infection of pathogens. Approximately 95% of these infections are associated with cancer cases, and recognized as infections of hepatitis B virus, hepatitis C virus, helicobacter pylori and human papilloma viruses. Consequently, deactivation of these pathogens signifies a promising therapy for this proportion of cancerous patients, suggesting that an important area in the future will be transcriptome and epigenome during tumorigenesis as well as manipulation of cancer genome for cancer modeling and therapy (Wen et al., 2016). CRISPR/Cas9-based genome modification is being explored in harnessing human immunodeficiency virus (HIV) infection by incorporating HIV into the host immune cells during its life cycle, where it works as a model for the expression of virus. Using genome manipulation technologies, investigators are currently testing 2 approaches to harness HIV infection. For the first technique, CRISPR/Cas9 nucleases target the viral genome sequence to eradicate the incorporated HIV genomic material from the genome of infected T cells (Liao et al., 2015). For the second technique, genome excision is used to enable HIV resistance by modifying chemokine receptor 5 (CCR5), a co-receptor required for HIV T cell infection. A scientist interrupted CCR5 in CD34⁺ hematopoietic stem and progenitor cells (HSPCs) by using the CRISPR/Cas9 genome editing technology and revealed 30% efficiency for the ablation of CCR5 in HSPCs CRISPR/Cas9 vectors (Mandal et al., 2014).

iii. CRISPR on the farm

With the extensive use of CRISPR, genome modification has become a cheap, fast and worthy tool in agriculture and livestock research and development (Ledford et al., 2015; Barrangou and Doudna, 2016). In the last decade, scientists have tried to create animals and plants with desired traits, suggesting an urgent need to design experiments or engineered pathways in plants to ensure the food security (Ledford et al., 2015). CRISPR could increase the profitability for farmers by reducing the need to cull animals. Biotechnologists are using the techniques to get desired gender from elite dairy and beef breeds to meet the need of milk and meat production and ensure food security (Reardon et al., 2016; Ahmad et al., 2017). The technique will facilitate the determination of the embryo sex at an early stage to cull the male embryos from layer hatch (Reardon et al., 2016; Tan et al., 2013).

Scientists have created polled animals that are being raised to spread genome through population. Cluster of Differentiation 163 (CD163) knockout pigs were generated by depleting CD163 complete protein or deletion of the region of *CD163* gene, encoding SRCR5 through using CRISPR/Cas9 gene target and somatic cell nuclear transfer (SCNT) approaches, which confers the resistance to infection of numerous porcine reproductive and respiratory syndrome (PRRSV) isolates in cultured cells or pigs. It was confirmed by correct folding and expression of SRCR5 removal *CD163* on macrophages surface and biological activity of hemoglobin hapto-globin scavenger (Yang et al., 2018; Burkard et al., 2017).

iv. Epigenetic modulation

CRISPR/Cas9 system has potential to remodel the epigenome at a specified locus by recruiting epigenetic modifiers through the sequence-specific DNA-binding proteins. It has been reported that the application of catalytically inactive CRISPR-associated 9 nuclease (dCas9) systems for site-specific epigenome editing with epigenetic modifiers induces alternation of epigenetic marks at their target DNA sites, resulting in relevant gene expression variation (Keung et al., 2014). A scientist (Kearns et al., 2015) targeted gene enhancers to retain the pluripotency in the embryonic stem cells of mouse by combining NmdCas9 with the histone demethylase LSD1, and found that NmdCas9-LSD1 could inhibit the expression of genes through targeted enhancers. The dCas9-LSD1 brings forth histone methylation for precision of gene expression repression (Kearns et al., 2015). Other researchers (Hilton et al., 2015) used a slightly different approach to fuse SpdCas9 and NmdCas9 to the histone acetyltransferase p300 catalytic core domain (SpdCas9-p300 core and Nm dCas9-p300 core, respectively). Such research is concentrated on targeting the enhancer or promoter regions to activate the expressions of endogenous genes. Given the wide range of functional epigenetic modulations from DNA methylation to histone modification, prospective studies are required to focus on the development of a complete toolkit of dCas9-based epigenetic modifiers (Figure 2B).

v. Live cell genomic imaging

In the post-genomic period, a contest for researchers is to recognize the relationship between the DNA genetic material and its structural arrangement in cell nucleus. Several studies have recognized that the structural arrangement of a genetic material may play a central role in the

regulatory gene expression and cell differentiation (Wang et al., 2016). The dCas9 fused to the fluorescent protein has been useful to picture particular genome loci in living cells. The event of telomeres tagged by dCas9-GFP or TRF1, a main telomere fusion protein, has been defined. This labeling approach provides a valuable tool for research of the dynamics of chromosome and structure and further extension of dCas9-based applications (Figure 2C) (Chen et al., 2013).

vi. Genome repression and activation

CRISPR/Cas9 was used as a powerful tool in both the gene therapy and biological research due to its efficiency of precise targeted genomic sequence within living cells and organisms (Ju et al., 2018). CRISPR/Cas9 has been preferred for a variety of novel functions, including expression of gene activation or repression (La Russa and Qi, 2015). For example, CRISPR/Cas9 Synergistic Activation Mediator (SAM) is a modified protein compound for the transcriptional stimulation of endogenous genes, which is based on the integration of MS2 RNA aptamers into loop two of the sgRNA and the stem-loop, which are exposed in the ternary Cas9 compound (Koneremann et al., 2015). Fusion of Cas9 with both light inducible and chemical systems resulted in an accurate spatiotemporal regulation of expression and activity at the transcriptional and post-translational levels. Doxycycline-inducible promoters accomplished inducible genome modifications in cells and mice by controlling the transcription of Cas9 and nCas9 (Dow et al., 2015). The post-translational control was investigated by insertion of intein-Cas9 combined proteins that renders Cas9 inactive. The intein sustain conformational modifications that activate protein splicing and reestablish Cas9 activity (Davis et al., 2015). In this system, an inducible genome modification at the post-translational level was achieved by introducing intein into Cas9 to restrict the nuclease activity (Figure 3). The dCas9 and effector components of CRISPRi/a gene regulating systems can also be combined with light inducible and chemical systems for spatiotemporal regulation. For instance, scientists (Polstein and Gersbach 2015) established the light inducible CRISPR activation systems to trigger gene expression endogenously. The improvement of split-Cas9 systems offers novel approaches for maintaining Cas9 activity, and it can expedite the delivery of Cas9 by circumventing the size restrictions of some transport systems (Figure 3). Besides Cas9, a DNase-dead Cpf1 mutant (ddCpf1) was generated for multiplex gene regulation. The difference of

dCas9 and ddCpf1 is that ddCpf1 could be used for gene repression in *E. coli*, and the repression was more active with crRNAs particularly directing the template filament of its target genes. When pointing the promoter area, both strands displayed active repression by the ddCpf1/crRNA complex. The high specificity of ddCpf1 enabled repression was further validated by the use of whole transcriptome RNA-seq techniques (Zhang et al., 2017).

vii. Base editing

Recent genome editing tools introduce DSB at a target locus as the first phase of gene manipulation (Cox et al., 2015). Although most of the genetic syndromes arise from point mutations, modern techniques for correcting point mutation are not efficient and normally induce random insertions and deletions at a target locus subsequent to cellular response to dsDNA breaks (Hilton and Gersbach 2015). However, the development of "base-editing" technique allows the direct, stable transformation of target DNA base into an alternative in a programmable way, without DNA double strand cleavage or a donor template (Komor et al., 2016). The fusion of CRISPR/Cas9 and cytidine deaminase enzyme was modified to facilitate the direct exchange of cytidine to uridine, thus producing a C→T (or G→A) substitution without induction of dsDNA breaks. The resultant "base-editors" transform cytidine to uridine within a frame of five nucleotides, and can proficiently correct a range of point mutations related to human syndrome (Kim et al., 2017). The DNA-base-editing in four human and murine cell lines used a Cas9 nickase targeting to fuse uracil glycosylase inhibitor in the second and third generation and manipulate the cellular DNA repair, resulting in a stable rectification of ~15-75% of total cellular DNA with least indel formation (Figure 4). Base editing has widened the scope and competency of genome editing of point mutations (Komor et al., 2016). APOBEC3B proteins were newly recognized as the mutational source in breast cancer cells that play roles in promoting survival of cells *in vivo*, by increasing DSB repair resulting genotoxic cures or spontaneous disruptions, therefore inhibiting cell death, and promoting a mutator phenotype, which regulates tumor progression. Hindering APOBEC3 catalytic activity and expression can increase sensitivity to genotoxic mediators and confines the development of mutations, which describes the cancer phenotype. These achievements can provide us with advanced model for cancer cell survival and tumor promotion and a mechanistic connection between ssDNA, DSBs, and grouped mutations.

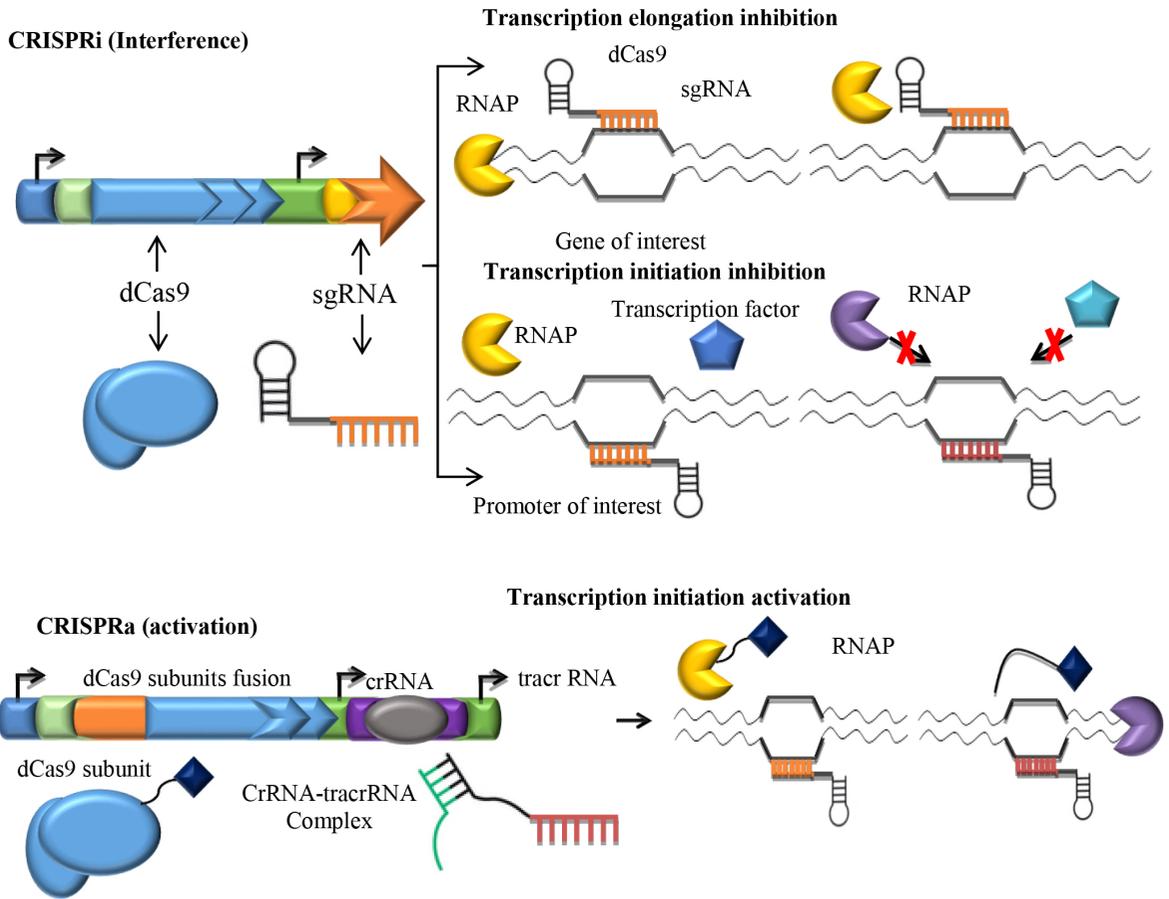


Figure 3. CRISPR interference and CRISPR activation. CRISPR interference system has been developed by modifying the type-II CRISPR/Cas system of *Streptococcus pyogenes* with a nuclease-dead Cas9 (dCas9) protein. Cas9 targeting uses a modified single guide RNA (sgRNA) that targets the dCas9 protein to particular DNA sequences that duplicate the CRISPR-RNA (crRNA) structure of the trans-acting CRISPR-RNA (tracrRNA) compound. This structure can be used for expression of a specific gene in two ways: binding dCas9-sgRNA composite to target DNA by preventing the transcription elongation, or binding the complex to the target DNA by stopping transcription initiation through hindering the RNA-polymerase (RNAP) binding or transcription factors. In both circumstances, transcription is suppressed. The similar structure can be used to trigger transcription by combining the RNAP- ω -subunit with dCas9. The ω -subunit is transported to the promoter by directing the dCas9-crRNA-tracrRNA complex to an upstream section of the gene, which enables transcription initiation by assembly of RNAP holoenzyme.

Therefore, anti-APOBEC3 agents could be favorable targets for cancers treatment in future (Lei et al., 2018).

viii. Engineered ecosystem

In addition, the use of CRISPR on organisms in the wild is also under consideration. The gene-drive technique could be used for rapid elimination of a modified gene from a population/rapid extension. The techniques can be used to eradicate different vectors with pathogens/diseases or plants prone to diseases and herbicide resistance (Ledford et al.,

2015). It takes time to spread genetic alteration from one organism to an entire population, because only half offspring can inherit a mutation on one pair of chromosomes (Barrangou and Doudna, 2016). However, the gene-drive technique along with CRISPR enables the chromosome to carry a mutation and copy itself for its partner in each generation, so that all the progenies will inherit the change (Gantz and Bier 2015; Ledford et al., 2015). It is generally believed that modifying or eradicating a whole population may have undesirable and unexpected effects on ecosystems and disturb the food chain.

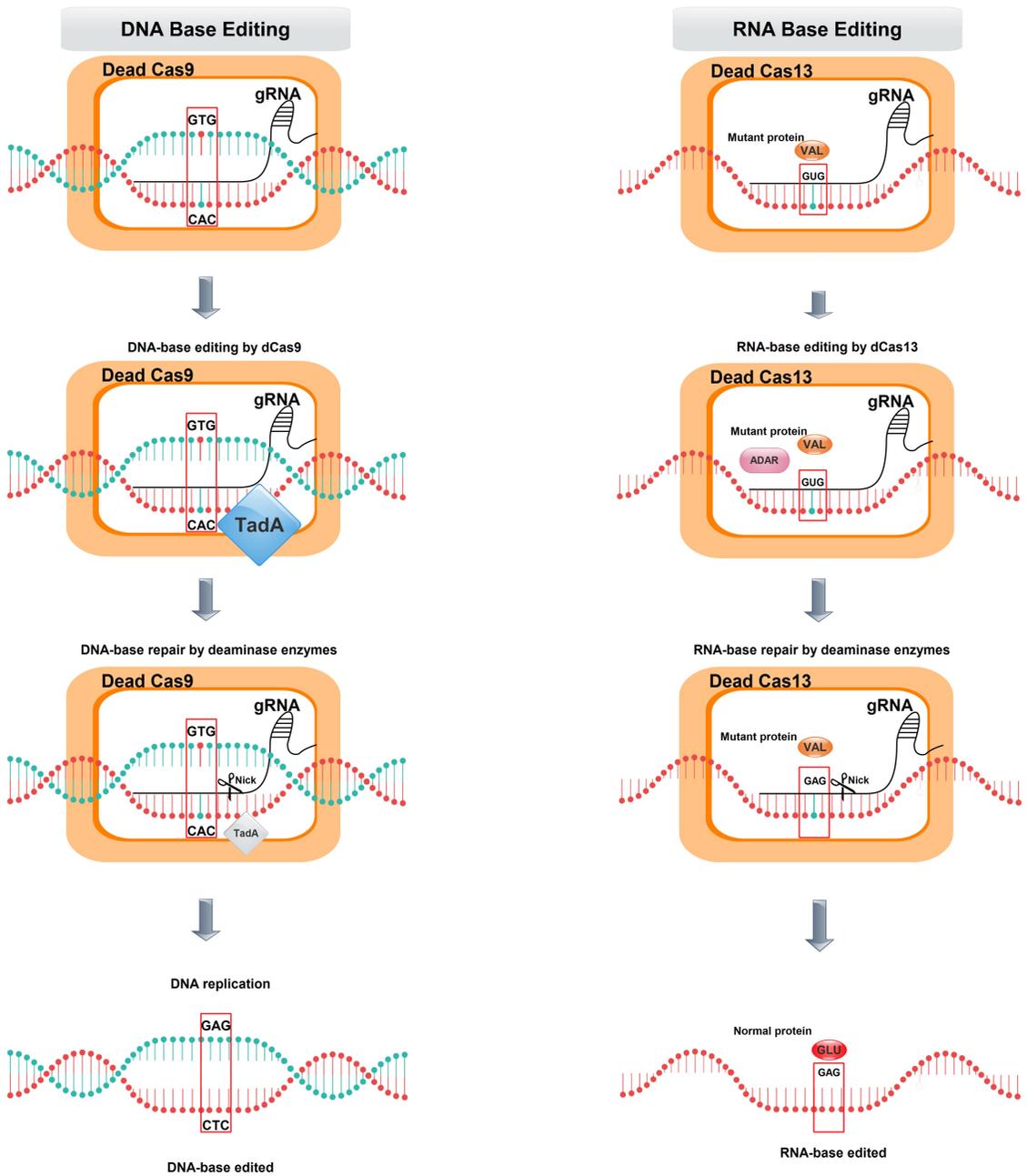


Figure 4. DNA-Base and RNA-Base editing.

DNA single base editing is interceded by generation of catalytically modified Cas9, deoxyadenosine deaminase enzymes and TadA fusion proteins complex. Resulting DNA recognition through single base editing-gRNA-target-DNA complex, a single stranded DNA is uncovered that comprises a targeted adenosine. Through deoxyadenosine-deaminase action, adenosine is transformed into inosine. Following DNA replication, or repair the primary targeted A-T base pair is replaced by G-C base pair.

RNA base editing is interceded by generation of catalytically modified dCas13b and ADAR complex, which naturally deaminates adenosine to inosine in RNA stand. The crRNA identifies the target site by hybridizing to the bases adjacent to the target adenosine, creating a RNA structure for editing, and recruiting the dCas13-ADAR fusion. A particular cytidine in the gRNA identifies a mismatched A for deamination to inosine, which functionally mimics guanosine in splicing and translation. VAL stands for valine, GLU stands for glutamate.

ix. Strategies of the delivery of the CRISPR/Cas components

To get the maximum precision and efficiency and output of CRISPR/Cas9 technologies, along with refining, the delivery systems used must be upgraded. The delivery approaches vary depending on application method and the cell type (Ahmad and Amiji, 2018). For instance, plasmids carrying Cas9 and sgRNA are conveyed through transitory transfection methods for small scale systematic application in cells, such as nucleofection, lipofectamine, and electroporation (Park et al., 2015; Lagutina et al., 2015), because incessant expression is not required once the preferred genome manipulation is achieved. However, transitory transfection might not be appropriate for repression analyses because CRISPRi only affects transcriptional efficiency relative to both protein and mRNA constancy (Carrol et al., 2016; Yang et al., 2016). Therefore, both non-viral and viral delivery methods with historic preferences in gene therapy are now being assessed for CRISPR/Cas9 delivery. Viral carriers that are possible delivery medium for CRISPR/Cas9 components contain self-deactivating lentivirus, and adenovirus. Vehicles for non-viral transport systems comprise plasmid DNA, Cas9 mRNA, in-vitro transcribed or produced sgRNA, Cas9/sgRNA ribonucleoprotein compounds (Patel et al., 2017), and donor nucleic acid prototypes. Researchers have several choices for delivering Cas9 and sgRNA into living cells. For example, cell lines can be transfected with lentivirus bearing encoded Cas9 and sgRNA, or these constituents can be delivered through electroporation (Yang et al., 2013). Though some syndromes can be treated through ex-vivo excision remedies, such as leukemia can be treated by transplanting edited bone marrow to the patient, most diseases will involve in-vivo transport of Cas9 and sgRNA (Glass et al., 2018).

Among viral transports, adeno virus based vectors are ideal because of their non-pathogenic features, minor immune reaction, which depends on the recognition of un-methylated CpG dinucleotides in the microbial DNA, and their aptitude to target non-dividing cells (Swiech et al., 2015). However, the bulky size of normal Cas9 confines the use of adeno virus based delivery methods, for which smaller 1-kb coding dCas9 was newly developed (Ran et al., 2015). For gene therapy, non-viral delivery techniques, such as hydrodynamic injections, have revealed encouraging results (Weber et al., 2015). Up until now, the suggested cellular delivery

preference is the use of purified Cas9-sgRNA ribonucleoproteins (RNPs) that permits the fast achievement of RNP composite in nucleus during its short stay inside the cells. RNP transport of Cas9 and sgRNA can be attained using various methods, including microinjection, electroporation, nucleotide transfection components, lipid mediated transfection, and cationic lipid mediated transfection reagents delivery systems (Ahmad and Amiji, 2018; Zuris et al., 2015). RNP delivery approaches are more specific and less damaging to cells as compared to other non-viral methods (Zuris et al., 2015). Several nano-particles, such as liposomes, lipids, inorganic nano-particles and nano-emulsions have revealed excellent outcomes as delivery carriers for transporting both drug proteins, and genes particularly in tumors because of their improved penetrability and holding effect (Schwank et al., 2013). Nano-particles are beneficial compared with other transport techniques because these improve the cellular uptake of nucleic acids, which are otherwise incapable to get inside the cell (Glass et al., 2018). Cationic lipid nano-particles have also recognized useful for gene and protein delivery. Intra-tumoral injection of Cas9 or sgRNA enclosed DNA nano-clews also revealed encouraging results in the cells, signifying the enhanced efficacy and accuracy of this method for genome edited protein delivery (Sun et al., 2015).

7. Potential challenges and recent advances in CRISPR/Cas systems

Despite being effectively useful in genome modification and functional regulation in biological systems, the CRISPR/Cas systems have also encountered various inherent potential challenges. The off-target effect is the foremost challenge. For instance, the off-target DSBs repair could result in insertions and deletions, leading to the occurrence of translocations and the activation of undesired consequences with other unknown off-target mutations. Off-target effects vary among different target DNA sequences due to variations in their genomic context. Different sgRNAs structure can affect the cleavage of on target and off targets sites. Potential off-target cleavage activity could occur on DNA sequence with even three to five base pair mismatches in the PAM-distal part of the sgRNA-guiding sequence. These mismatches of the one to five base pairs at the 5' end of sgRNAs are better tolerated than those at the 3' end (Zhang et al., 2015). Target recognition *in vitro* and in bacterial cells requires the "seed" sequence at the 3' end of 1-7 bp adjacent to PAM (Jiang et al., 2013; Jinek et al., 2012).

In addition, within the targeting region of sgRNA, the effects of location-based mismatches on CRISPR/Cas9 activity are not predictable, probably striking for mismatches at 5' end and negligible for those at 3' end. Presumptive solutions to this problem include the use of bioinformatics tool for selection of sgRNAs (www.biotoools.com), such as CRISPR-offinder (Zhao et al., 2017) and direct use of inducible and non-inducible CRISPR/Cas9 proteins (Müller et al., 2016; Slaymaker et al., 2016). These applications have had partial achievement due to the different sequences and genetic framework of each target DNA. Illustration of the interaction mechanisms of sgRNA-target DNA unlocks the gateway for the rational scheme of CRISPR/Cas9 along with enhanced specificities as validated by both SpCas9-HF1 (Cas9 from *Streptococcus pyogenes*) and SpCas9 with dissimilar arrangements of amino acid alterations (Slaymaker et al., 2016). Therefore, the position of the protospacer sequence in each base pair is central to the identified amino acids for targeted engineering of the CRISPR/Cas9 tailored to all the target DNA sites.

Reported strategies for increased Cas9 specificity include reduction of dynamic Cas9 within the cell by using Cas9 nickase mutants to generate a pair of juxtaposed single-stranded DNA nicks, and truncation of the guide sequence at the 5' end by fusing a pair of catalytically inactive Cas9 nucleases to a *FokI* nuclease domain. Although they decrease off-target mutagenesis, the two strategies have a number of restrictions: On-target cleavage efficiency can reduce with the decrease of the amount of Cas9 nucleases; double nicking requires the simultaneous transport of two sgRNAs; and truncated guides can induce indel formation at some off-target loci and lessen the number of target sites in the genome (Slaymaker et al., 2016). The SpCas9 crystal structure in composite with sgRNA and target DNA offers a basis to develop specificity by rational engineering. The structure discloses a positively charged groove, located among the interacting domains of the RuvC, HNH, and PAM in SpCas9, which is involved in stabilizing the non-target strand of target DNA (Jinek et al., 2012). Unbiased whole-genome off-target analysis and targeted deep sequencing were used to assess Cas9-mediated DNA cleavage in human cells (Kuscu et al., 2014). It can be concluded that enhanced specificity in SpCas9 (eSpCas9) variants has diminished the off-target effects as well as increased durable on-target cleavage efficiency, enabling the applications of

eSpCas9 in genome editing which requires high specificity (Li et al., 2014).

i. RNA editing

Genome editing techniques hold promises for the treatment of genetic diseases at the RNA level by blocking the disease-associated sequences from producing disease-causing proteins (Hannon and Rossi, 2004). Type VI CRISPR/Cas systems possess Cas13 containing single effector RNA guided RNases. The Cas13 ortholog from Type VI system was engineered to establish RNA editing and knockdown by using catalytically inactive Cas13 (dCas13) to switch on the adenosine-to-inosine deaminase activity in mammalian cells. This system has been declared as RNA Editing for Programmable A to I Replacement (REPAIR) with the advantages of no sequence restrictions and the ability to correct full-length transcripts containing pathogenic mutations (Figure 4). Furthermore, this system can also be used to produce a high specificity variant and reduce the system for viral delivery. REPAIR offers a favorable RNA editing platform with broad applicability for research, therapeutics, and biotechnology (Kim et al., 2018; Cox et al., 2017).

ii. Anti-CRISPR/Cas system

Active inhibitors of CRISPR/Cas systems were unveiled in phages of the closely-linked group of *Pseudomonas* spp. Phages with the active type I-F CRISPR/Cas system were capable to infect and proliferate in a *Pseudomonas aeruginosa* strain even with the proto-spacer system that should have been directed by this system (Bondy-Denomy et al., 2013). Anti-CRISPR genomic regions were identified by analysis of sequences that are responsible for this phenotype. A total of five distinctive proteins (AcrF1-AcrF5) were exposed to deactivate the type I-F CRISPR/Cas system. Up till now, there have been twenty one typical groups of anti-CRISPR proteins defined against type I and type II systems. Moreover, for the proteins with a size ranging 50-150 amino acids, there are no shared structures that are conserved between their sequences. None of groups has sequence resemblance to any protein of predicted functions. Over the past few years, action mechanisms have been revealed for 6 anti-CRISPR proteins by using a combination of biochemical, structural genetic, and studies (Pawluk et al., 2017).

8. Conclusions and future perspectives

CRISPR/Cas technology would not only have potential therapeutic applications, but also it is one

of an invaluable molecular tool in research. CRISPR's talent to accurately modify DNA sequences makes for further specific amendments, but it makes it more challenging for researchers to recognize a genetically modified organism once it has been engineered. With genome editing, there is no longer talent to truly track the engineered products, and it may be difficult to distinguish whether something has been changed genetically or engineered conventionally. However, the main interest in CRISPR/Cas9 advancement has been mainly to merge all the strategies into one and manufacture "super" version of CRISPR/Cas9, apt for the entire target DNA, the desired aim that may not be persuasively achievable. Future goals are aimed to evaluate the effect of anti-CRISPR genes in horizontal gene transfer, their influence on virulence genes and antibiotic resistance spread is one of the major goals.

Acknowledgements

This work was supported by the National Key Research and Development Program of China, Stem Cell and Translational Research [2016YFA0100203], the National Transgenic Project of China [2016ZX08006003-004], and the Fundamental Research Funds for the Central Universities [2662015BQ005 and 2662016PY006]. We thank Hanchang Zhu for the linguistic assistance during the preparation of this manuscript.

References

Ahmad, G., and Amiji, M. (2018). Use of CRISPR/Cas9 gene-editing tools for developing models in drug discovery. *Drug discovery today*. <https://doi.org/10.1016/j.drudis.2018.01.014>.

Ahmad, H.I., Yasir, M., Ahmad, M.J., (2017). CRISPR/Cas9 system: A broadly applicable tool for targeted genome engineering. *J Innov Bio-Res.* 1, 13-23.

Amitai, G., and Sorek, R. (2016). CRISPR-Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.* 14, 67-76. doi:10.1038/nrmicro.2015.14.

Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., and Pul, Ü. (2014). Detection and characterization of spacer integration intermediates in type I E CRISPR-Cas system. *Nucleic acid Res.* 42, 7884-7893. <https://doi.org/10.1093/nar/gku510>.

Barrangou, R., Fremaux, C., and Deveau, H. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709-1712. DOI: 10.1126/science.1138140.

BARRANGOU, R. and DOUDNA, J. A. (2016). Applications of CRISPR technologies in research

and beyond. *Nature biotechnology*, 34, 933. doi: 10.1038/nbt.3659

Beerli, RR., and Barbas, CF. (2002). Engineering polydactyl zinc-finger transcription factors. *Nat. Biotechnol.* 20, 135-141. doi:10.1038/nbt0202-135.

Bhakta, MS., Henry, IM., and Ousterout, DG. (2013). Highly active zinc-finger nucleases by extended modular assembly. *Genome Research* 23, 530-538. doi/10.1101/gr.143693.112.

Bibikova, M., Beumer, K., Trautman, JK., and Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. *Science* 300, 764-764. DOI: 10.1126/science.1079512.

Bibikova, M., Golic, M., Golic, KG., and Carroll, D. (2002). Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* 161, 1169-1175.

Bickle, TA., and Krüger, D. (1993). Biology of DNA restriction. *Microbiol. Rev.* 57, 434-450.

Bingham, R., Ekunwe, SI., Falk, S., Snyder, L., and Kleanthous, C. (2000). The major head protein of bacteriophage T4 binds specifically to elongation factor Tu. *J. Biol. Chem.* 275, 23219-23226. doi: 10.1074/jbc.M002546200.

Boch, J., Scholze, H., and Schornack, S. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509-1512. DOI: 10.1126/science.1178811.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, SD. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology.* 151, 2551-2561. doi: 10.1099/mic.0.28048-0.

Bondy-Denomy, J., Pawluk, A., Maxwell, KL., and Davidson, AR. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature.* 493, 429-432. doi:10.1038/nature11723.

Bosch, Mvd., Lohman, PH., and Pastink, A. (2002). DNA double-strand break repair by homologous recombination. *Biol. Chem.* 383, 873-892. <https://doi.org/10.1515/BC.2002.095>.

BURKARD, C., LILLICO, S. G., REID, E., JACKSON, B., MILEHAM, A. J., AIT-ALI, T., WHITELAW, C. B. A. and ARCHIBALD, A. L. (2017). Precision engineering for PRRSV resistance in pigs: Macrophages from genome edited pigs lacking CD163 SRCR5 domain are fully resistant to both PRRSV genotypes while maintaining biological function. *PLoS pathogens*, 13, e1006206. <https://doi.org/10.1371/journal.ppat.1006206>

Burnet, F. (1929). Further observations on the nature of bacterial resistance to bacteriophage. *J.*

- Pathol.* 32, 349-354. DOI: 10.1002/path.1700320302.
- Carroll, K. J., Makarewich, C. A., McAnally, J., Anderson, D. M., Zentilin, L., Liu, N., ... and Olson, E. N. (2016). A mouse model for adult cardiac-specific gene deletion with CRISPR/Cas9. *Proceedings of the National Academy of Sciences*, 113, 338-343. <https://doi.org/10.1073/pnas.1523918113>.
- Carthew, RW., and Sontheimer, EJ. (2009). Origins and mechanisms of miRNAs and siRNAs. *Cell* 136, 642-655. <https://doi.org/10.1016/j.cell.2009.01.035>.
- Cermak, T., Doyle, EL., and Christian, M. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids research*, gkr218. <https://doi.org/10.1093/nar/gkr218>.
- Chen, B., Gilbert, LA., and Cimini, BA. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155, 1479-1491. <https://doi.org/10.1016/j.cell.2013.12.001>.
- CHEN, J. S., DAGDAS, Y. S., KLEINSTIVER, B. P., WELCH, M. M., SOUSA, A. A., HARRINGTON, L. B., STERNBERG, S. H., JOUNG, J. K., YILDIZ, A. and DOUDNA, J. A. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*, 550, 407. doi:10.1038/nature24268.
- Childs, LM., England, WE., Young, MJ., Weitz, JS., and Whitaker, RJ. (2014). CRISPR-induced distributed immunity in microbial populations. *PLoS One* 9, e101710. <https://doi.org/10.1371/journal.pone.0101710>.
- Cho, J., Parks, ME., and Dervan, PB. (1995). Cyclic polyamides for recognition in the minor groove of DNA. *P. Nat. Acad. Sci.* 92, 10389-10392.
- Christian, M., Cermak, T., and Doyle, EL. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. 186, 757-761. <https://doi.org/10.1534/genetics.110.120717>.
- Chu, VT., Weber, T., and Wefers, B. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nature biotechnology* 33, 543-548. doi:10.1038/nbt.3198.
- Chylinski, K., Makarova, KS., Charpentier, E., and Koonin, EV. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research* 42, 6091-6105. <https://doi.org/10.1093/nar/gku241>.
- Cong, L., Ran, FA., and Cox, D. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823. DOI: 10.1126/science.1231143.
- Cox, DB., Gootenberg, JS., and Abudayyeh OO. (2017). RNA editing with CRISPR-Cas13. *Science*, 358, 1019-1027. DOI: 10.1126/science.aaq0180.
- Cox, DBT., Platt, RJ., and Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. *Nature medicine* 21, 121-131. doi: 10.1038/nm.3793.
- d'Herelle, F. (1917). Sur un microbe invisible antagoniste des bacilles dysentériques. *CR Acad. Sci. Paris* 165, 373-375.
- Datsenko, KA., Pougach, K., and Tikhonov, A. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* 3, 945. doi: 10.1038/ncomms1937.
- Davis, KM., Pattanayak, V., Thompson, DB., Zuris, JA., and Liu, DR. (2015). Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat. Chem. Biol.* 11, 316-318. doi:10.1038/nchembio.1793.
- Deltcheva, E., Chylinski, K., and Sharma, CM. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. 471, 602-607. doi:10.1038/nature09886.
- Deng, D., Yan, C., and Pan, X. (2012). Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science*. 335, 720-723. DOI: 10.1126/science.1215670.
- Deveau, H., Barrangou, R., and Garneau, JE. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390-1400. doi: 10.1128/JB.01412-07.
- Doudna, JA., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*. 346, 1258096. DOI: 10.1126/science.1258096.
- Dow, LE. (2015). Modeling disease in vivo with CRISPR/Cas9. *Trends. Mol. Medicine*. 21, 609-621. <https://doi.org/10.1016/j.molmed.2015.07.006>.
- Dow, LE., Fisher, J., and O'rourke, KP. (2015). Inducible in vivo genome editing with CRISPR-Cas9. *Nat. biotechnol.* 33, 390-394. doi:10.1038/nbt.3155.
- Dow, LE., and Lowe, SW. (2012). Life in the fast lane: mammalian disease models in the genomics era. *Cell*. 148, 1099-1109. <https://doi.org/10.1016/j.cell.2012.02.023>.
- Faruqi, AF., Egholm, M., and Glazer, PM. (1998). Peptide nucleic acid-targeted mutagenesis of a

- chromosomal gene in mouse cells. *P. Nat. Acad. Sci.* 95, 1398-1403.
- Faruqi, AF., Seidman, MM., Segal, DJ., Carroll, D., and Glazer, PM. (1996). Recombination induced by triple-helix-targeted DNA damage in mammalian cells. *Mol. Cell. Biol.* 16, 6820-6828.
- Fineran, PC., and Charpentier, E. (2012). Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information. *Virology* 434, 202-209. <https://doi.org/10.1016/j.virol.2012.10.003>.
- Fineran, PC., Gerritzen, MJ., and Suárez-Diez, M. (2014). Degenerate target sites mediate rapid primed CRISPR adaptation. *P. Nat. Acad. Sci.* 111, E1629-E1638. doi: 10.1073/pnas.1400071111.
- Friedland, AE., Tzur, YB., and Esvelt, KM., (2013). Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nature Methods*. 10, 741-743. doi:10.1038/nmeth.2532.
- Gaj, T., Gersbach, CA., and Barbas, CF. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trend. Biotechnol.* 31, 397-405. <https://doi.org/10.1016/j.tibtech.2013.04.004>.
- Gantz, VM., and Bier, E. (2015). The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science*. 348, 442-444. DOI: 10.1126/science.aaa5945.
- Glass, Z., Lee, M., Li, Y., and Xu, Q. (2018). Engineering the Delivery System for CRISPR-Based Genome Editing. *Trends in biotechnology*. 36, 173-185. <https://doi.org/10.1016/j.tibtech.2017.11.006>.
- Gonzalez, B., Schwimmer, LJ., and Fuller, RP. (2010). Modular system for the construction of zinc-finger libraries and proteins. *Nature Protocols*. 5, 791-810. doi:10.1038/nprot.2010.34.
- Gottesfeld, JM., Neely, L., Trauger, JW., Baird, EE., and Dervan, PB. (1997). Regulation of gene expression by small molecules. *Nature*. 387, 202.
- Haft, DH., Selengut, J., Mongodin, EF., and Nelson, KE. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 1, e60. <https://doi.org/10.1371/journal.pcbi.0010060>.
- Hanlon, GW., Denyer, SP., Olliff, CJ., and Ibrahim, LJ. (2001). Reduction in exopolysaccharide viscosity as an aid to bacteriophage penetration through *Pseudomonas aeruginosa* biofilms. *Appl. Environ. Microbiol.* 67, 2746-2753. doi: 10.1128/AEM.67.6.2746-2753.2001.
- Hannon, G. J., and Rossi, J. J. (2004). Unlocking the potential of the human genome with RNA interference. *Nature*. 431, 371.
- Hilton, IB., D'ippolito, AM., and Vockley, CM. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33, 510-517. doi: 10.1038/nbt.3199.
- Hilton, IB., and Gersbach, CA. (2015). Enabling functional genomics with genome engineering. *Genome research* 25, 1442-1455. doi:10.1101/gr.190124.115
- Hockemeyer, D., Wang, H., and Kiani, S. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. biotechnol.* 29, 731-734. doi:10.1038/nbt.1927.
- Hooton, S., and Connerton, IF. (2015). *Campylobacter jejuni* acquire new host-derived CRISPR spacers when in association with bacteriophages harboring a CRISPR-like Cas4 protein. *Frontiers in Microbiology* 5, 744. <https://doi.org/10.3389/fmicb.2014.00744>.
- Hruscha, A., Krawitz, P., and Rechenberg, A. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development*. 140, 4982-4987. doi: 10.1242/dev.099085.
- 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. doi: 10.1128/jb.169.12.5429-5433.1987.
- Jansen, R., Embden, J., Gaastra, W., and Schouls, L. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565-1575. DOI: 10.1046/j.1365-2958.2002.02839.
- Jiang, W., Bikard, D., Cox, D., Zhang, F., Marraffini, LA. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233-239. doi:10.1038/nbt.2508.
- Jinek, M., Chylinski, K., and Fonfara, I. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821. DOI: 10.1126/science.1225829.
- Jinek, M., East, A., and Cheng, A. (2013). RNA-programmed genome editing in human cells. *elife* 2, e00471. DOI: 10.7554/eLife.00471.001.
- Jung, T-Y., An, Y., and Park, K-H. (2015). Crystal structure of the Csm1 subunit of the Csm complex and its single-stranded DNA-specific nuclease

- activity. *Structure*. 23, 782-790. DOI:10.7554/eLife.00471.001.
- JU, X. D., XU, J. and SUN, Z. S. 2018. CRISPR Editing in Biological and Biomedical Investigation. *Journal of cellular biochemistry*. 119, 52-61. DOI 10.1002/jcb.26154
- Kapitonov, VV., Makarova, KS., and Koonin, EV. (2016). ISC, a novel group of bacterial and archaeal DNA transposons that encode Cas9 homologs. *J. Bacteriol.* 198, 797-807. doi: 10.1128/JB.00783-15.
- Kearns, NA., Pham, H., Tabak, B. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nature Methods*. 12, 401-403. doi:10.1038/nmeth.3325.
- Keung, AJ., Bashor, CJ., Kiriakov, S., Collins, JJ., Khalil, AS. (2014). Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. *Cell*. 158, 110-120. <https://doi.org/10.1016/j.cell.2014.04.047>.
- Kim, H., and Kim, J-S. (2014). A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* 15, 321-334. doi:10.1038/nrg3686.
- Kim, S., Lee, MJ., Kim, H., Kang, M., and Kim, J-S. (2011). Preassembled zinc-finger arrays for rapid construction of ZFNs. *Nature Methods*. 8, 7.
- Kim, V. N. (2018). RNA-targeting CRISPR comes of age. *Nature biotechnology*, 36, 44.
- Kim, Y. B., Komor, A. C., Levy, J. M., Packer, M. S., Zhao, K. T., and Liu, D. R. (2017). Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nature biotechnology*, 35, 371. doi: 10.1038/nbt.380
- Kim, Y-G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *P. Nat. Acad. Sci.* 93, 1156-1160.
- Komor, AC., Kim, YB., Packer, MS., Zuris, JA., and Liu, DR. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 533, 420-424. doi: 10.1038/nature17946.
- Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., ... and Nureki, O. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 517(7536), 583. doi:10.1038/nature14136
- Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature biotechnology*, 32, 677-683. doi:10.1038/nbt.2916
- Lagutina, I. V., Valentine, V., Picchione, F., Harwood, F., Valentine, M. B., Villarejo-Balcells, B., ... and Grosveld, G. C. (2015). Modeling of the human alveolar rhabdomyosarcoma Pax3-Foxo1 chromosome translocation in mouse myoblasts using CRISPR-Cas9 nuclease. *PLoS genetics*, 11(2), e1004951. <https://doi.org/10.1371/journal.pgen.1004951>.
- LA RUSSA, M. F. and QI, L. S. 2015. The new state of the art: Cas9 for gene activation and repression. *Molecular and cellular biology*, 35, 3800-3809. doi: 10.1128/MCB.00512-15
- Ledford, H. (2015). CRISPR, the disruptor. *Nature*. 522, 20.
- Lee, KH., Lee, SG., Eun, Lee K. (2012). Identification, structural, and biochemical characterization of a group of large Csn2 proteins involved in CRISPR-mediated bacterial immunity. *Proteins: Structure, Function, and Bioinformatics*. 80, 2573-2582. DOI: 10.1002/prot.24138.
- Levy, A., Goren, MG., Yosef, I. (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*. 520, 505-510. doi:10.1038/nature14302.
- Lei, L., Chen, H., Xue, W., Yang, B., Hu, B., Wei, J., ... and Yan, L. (2018). APOBEC3 induces mutations during repair of CRISPR-Cas9-generated DNA breaks. *Nature structural and molecular biology*, 25, 45. doi:10.1038/s41594-017-0004-6
- Li, M., Wang, R., Zhao, D., and Xiang, H. (2014). Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. *Nucleic acids research*. 42, 2483-2492. <https://doi.org/10.1093/nar/gkt1154>.
- Liao, H-K., Gu, Y., and Diaz, A. (2015). Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nature communications*. 6. doi:10.1038/ncomms7413
- Lin, S-R., Yang, H-C., and Kuo, Y-T. (2014). The CRISPR/Cas9 system facilitates clearance of the intrahepatic HBV templates in vivo. *Molecular Therapy-Nucleic Acids* 3, e186. <https://doi.org/10.1038/mtna.2014.38>.
- Liu, Q., Segal, DJ., Ghiara, JB., and Barbas, CF. (1997). Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *P. Nat. Acad. Sci.* 94, 5525-5530.
- Luria, SE., and Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*. 28, 491.
- Makarova, KS., Anantharaman, V., Aravind, L., and Koonin, EV. (2012). Live virus-free or die: coupling of antiviral immunity and programmed suicide or

- dormancy in prokaryotes. *Biology Direct*. 7, 40. <https://doi.org/10.1186/1745-6150-7-40>.
- Makarova, KS., Anantharaman, V., Grishin, NV., Koonin, EV., and Aravind, L. (2014). CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *F. gene*. 5, 102. doi: 10.3389/fgene.2014.00102
- Makarova, KS., Aravind, L., Grishin, NV., Rogozin, IB., and Koonin, EV. (2002). A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research*. 30, 482-496. <https://doi.org/10.1093/nar/30.2.482>.
- Makarova, KS., Aravind, L., Wolf, YI., and Koonin, EV. (2011). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biology Direct*. 6, 38. <https://doi.org/10.1186/1745-6150-6-38>.
- Makarova, KS., Grishin, NV., Shabalina, SA., Wolf, YI., and Koonin, EV. (2006). 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, 7. <https://doi.org/10.1186/1745-6150-1-7>.
- MAKAROVA, K. S., HAFT, D. H., BARRANGOU, R., BROUNS, S. J., CHARPENTIER, E., HORVATH, P., MOINEAU, S., MOJICA, F. J., WOLF, Y. I. and YAKUNIN, A. F. (2011). Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology*, 9, 467. doi:10.1038/nrmicro2577
- Makarova, KS., Wolf, YI., and Alkhnbashi, OS. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722-736. doi:10.1038/nrmicro3569.
- Makarova, KS., Wolf, YI., and Koonin, EV. (2013). The basic building blocks and evolution of CRISPR-Cas systems. Portland Press Limited, pp. 1392-1400. DOI: 10.1042/BST20130038.
- Mali, P., Esvelt, KM., and Church, GM. (2013a). Cas9 as a versatile tool for engineering biology. *Nature Methods*. 10, 957-963. doi:10.1038/nmeth.2649.
- Mali, P., Yang, L., and Esvelt, KM. (2013b). RNA-guided human genome engineering via Cas9. *Science*. 339, 823-826. DOI: 10.1126/science.1232033.
- Mandal, PK., Ferreira, LM., and Collins R. (2014). Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell stem cell*. 15, 643-652. <https://doi.org/10.1016/j.stem.2014.10.004>.
- Marraffini, LA., and Sontheimer, EJ. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843-1845. DOI: 10.1126/science.1165771.
- Maruyama, T., Dougan, SK., Truttmann, MC. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33, 538-542. doi:10.1038/nbt.3190.
- Mercer, AC., Gaj, T., Fuller, RP., and Barbas, CF. (2012). Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Research*. 40, 11163-11172. <https://doi.org/10.1093/nar/gks875>.
- Miller, JC., Tan, S., and Qiao, G. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* 29, 143-148. doi:10.1038/nbt.1755.
- Mohanraju, P., Makarova, KS., and Zetsche, B. (2016). Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science*. 353, aad5147. DOI: 10.1126/science.aad5147.
- Mojica, F., Díez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 155, 733-740. doi: 10.1099/mic.0.023960-0.
- Mojica, F.J., Díez-Villaseñor, C., Soria, E., Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* 36, 244-246.
- Mojica, F.J., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174-182. DOI: 10.1046/j.1365-2958.2000.01838.x.
- Moscou, MJ., and Bogdanove, AJ. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science*. 326, 1501-1501. DOI: 10.1126/science.1178817.
- Mulepati, S., Héroux, A., and Bailey, S. (2014). Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science*. 345, 1479-1484. DOI: 10.1126/science.1256996.
- Müller, M., Lee, CM., and Gasiunas, G. (2016). *Streptococcus thermophilus* CRISPR-Cas9 systems enable specific editing of the human genome. *Molecular Therapy*. 24, 636-644. <https://doi.org/10.1038/mt.2015.218>.
- Mussolino, C., Morbitzer, R., Lütge, F. (2011). A novel TALE nuclease scaffold enables high

- genome editing activity in combination with low toxicity. *Nucleic Acids Research*. 39, 9283-9293. <https://doi.org/10.1093/nar/gkr597>.
- Nuñez, JK., Kranzusch, P.J., and Noeske, J. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* 21, 528-534. doi: 10.1038/nsmb.2820.
- Nuñez, JK., Lee, AS., Engelman, A., and Doudna, JA. (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature*. 519, 193-198. doi:10.1038/nature14237.
- Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015). Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Molecular Cell*. 58, 418-430. <https://doi.org/10.1016/j.molcel.2015.03.018>.
- Paez-Espino, D., Sharon, I., and Morovic W. (2015). CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *MBio*. 6, e00262-00215. doi: 10.1128/mBio.00262-15.
- Park, C. Y., Kim, D. H., Son, J. S., Sung, J. J., Lee, J., Bae, S., ... and Kim, J. S. (2015). Functional correction of large factor VIII gene chromosomal inversions in hemophilia A patient-derived iPSCs using CRISPR-Cas9. *Cell stem cell*, 17, 213-220. <https://doi.org/10.1016/j.stem.2015.07.001>.
- Parma, DH., Snyder, M., and Sobolevski, S. (1992). The Rex system of bacteriophage lambda: tolerance and altruistic cell death. *Genes and Development*. 6, 497-510.
- Pavletich, NP., Pabo, CO. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*. 252, 809.
- Pawluk, A., Davidson, AR., and Maxwell, KL. (2017). Anti-CRISPR: discovery, mechanism and function. *Nature Reviews Microbiology*. nrmicro.2017.2120.
- Polstein, LR., and Gersbach, CA. (2015). A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* 11, 198-200. doi:10.1038/nchembio.1753.
- 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.
- Ran, FA., Hsu, PD., Wright, J. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*. 8, 2281-2308. DOI:10.1038/nprot.2013.143.
- Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., ... and Koonin, E. V. (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*, 520, 186-191. doi:10.1038/nature14299
- Reardon, S. (2016). The CRISPR zoo. *Nature*. 531, 160-163.
- Rouillon, C., Zhou, M., and Zhang, J. (2013). Structure of the CRISPR interference complex CSM reveals key similarities with cascade. *Molecular Cell*. 52, 124-134. doi: 10.1016/j.molcel.2013.08.020.
- Sander, JD., and Joung, JK. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347-355. doi: 10.1038/nbt.2842.
- Sandor, Z., and Bredberg, A. (1995). Deficient DNA repair of triple helix-directed double psoralen damage in human cells. *FEBS Letters*. 374, 287-291. [https://doi.org/10.1016/0014-5793\(95\)01133-Y](https://doi.org/10.1016/0014-5793(95)01133-Y).
- Sanjana, NE., Cong, L., and Zhou, Y. (2012). A transcription activator-like effector toolbox for genome engineering. *Nature Protocols*. 7, 171-192. doi: 10.1038/nprot.2011.431.
- Schwank, G., Koo, B. K., Sasselli, V., Dekkers, J. F., Heo, I., Demircan, T., ... and Nieuwenhuis, E. E. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell stem cell*, 13, 653-658. <https://doi.org/10.1016/j.stem.2013.11.002>.
- Shmakov, S., Abudayyeh, OO., and Makarova, KS. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Molecular Cell*. 60, 385-397. doi:10.1016/j.molcel.2015.10.008.
- Silas, S., Mohr, G., and Sidote, DJ. (2016). Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science*. 351, aad4234. doi: 10.1126/science.aad4234.
- Slymaker, IM., Gao, L., and Zetsche, B. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science*. 351, 84-88. doi: 10.1126/science.aad5227.
- Spilman, M., Cocozaki, A., and Hale, C. (2013). Structure of an RNA silencing complex of the CRISPR-Cas immune system. *Molecular Cell*. 52, 146-152. doi: 10.1016/j.molcel.2013.09.008.
- Strobel, SA., and Dervan, PB. (1990). Site-specific cleavage of a yeast chromosome by oligonucleotide-directed triple-helix formation. *Science*. 249, 73-76.
- Strobel, SA., Dervan, PB. (1991). Single-site enzymatic cleavage of yeast genomic DNA mediated by triple helix formation. *Nature*. 350, 172. doi:10.1038/350172a0.

- Sun, W. et al. (2015) Self-assembled DNA nanoclews for the efficient delivery of CRISPR-Cas9 for genome editing. *Angew. Chem. Int. Ed. Engl.* 54, 12029-12033.
- Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., ... and Zhang, F. (2015). In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nature biotechnology*, 33, 102-106. doi:10.1038/nbt.3055.
- Tan, W., Carlson, DF., and Lancto, CA. (2013). Efficient nonmeiotic allele introgression in livestock using custom endonucleases. *P. Nat. Acad. Sci.* 110, 16526-16531. doi: 10.1073/pnas.1310478110.
- Taylor, DW., Zhu, Y., and Staals, RH. (2015). Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science*. 348, 581-585.
- Terns, MP., and Terns, RM. (2011). CRISPR-based adaptive immune systems. *Current Opin. Microbiol.* 14, 321-327. doi: 10.1126/science.aaa4535.
- Thomas, KR., and Capecchi, MR. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell*. 51, 503-512.
- Twort, FW. (1915). An investigation on the nature of ultra-microscopic viruses. *The Lancet*. 186, 1241-1243. [http://dx.doi.org/10.1016/S0140-6736\(01\)20383-3](http://dx.doi.org/10.1016/S0140-6736(01)20383-3).
- Urnov, FD., Rebar, EJ., Holmes, MC., Zhang, HS., and Gregory, PD. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11, 636-646. doi: 10.1038/nrg2842.
- Varshavsky, A. (2006). Discovering the RNA double helix and hybridization. *Cell*. 127, 1295-1297. DOI: 10.1016/j.cell.2006.12.008.
- Vestergaard, G., Garrett, RA., and Shah, SA. (2014). CRISPR adaptive immune systems of Archaea. *RNA Biology*. 11, 156-167. doi: 10.4161/ma.27990.
- Wang, H., La, Russa, M., and Qi, LS. (2016). CRISPR/Cas9 in genome editing and beyond. *Annual Review of Biochemistry*. 85, 227-264. doi: 10.1146/annurev-biochem-060815-014607.
- Wang R, Preamplume G, Terns MP, Terns RM, Li H (2011) Interaction of the Cas6 ribonuclease with CRISPR RNAs: recognition and cleavage. *Structure* 19, 257-264. doi: 10.1016/j.str.2010.11.014.
- Weber, J., Öllinger, R., Friedrich, M., Ehmer, U., Barenboim, M., Steiger, K., ... and Gross, N. (2015). CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. *Proceedings of the National Academy of Sciences*, 112, 13982-13987. <https://doi.org/10.1073/pnas.1512392112>
- Wei, Y., Terns, RM., and Terns, MP. (2015). Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes and Development*. 29, 356-361. doi:10.1101/gad.257550.114.
- Weinberger, AD., Sun, CL., and Pluciński, MM. (2012) Persisting viral sequences shape microbial CRISPR-based immunity. *PLoS Computat. Biol.* 8, e1002475.
- Wen, WS., Yuan, ZM., Ma, SJ., Xu, J., and Yuan, DT. (2016). CRISPR-Cas9 systems: versatile cancer modelling platforms and promising therapeutic strategies. *Int. J. Cancer*. 138, 1328-1336. doi: 10.1002/ijc.29626.
- Wood, AJ., Lo, T-W., and Zeitler, B. (2011). Targeted genome editing across species using ZFNs and TALENs. *Science*. 333, 307-307. doi: 10.1126/science.1207773.
- Wyman, C., and Kanaar, R. (2006). DNA double-strand break repair: all's well that ends well. *Annual Review of Genetics*. 40, 363-383. DOI: 10.1146/annurev.genet.40.110405.090451.
- Xie, F., Ye, L., and Chang, JC., . (2014). Seamless gene correction of β -thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Research*. 24, 1526-1533. doi: 10.1101/gr.173427.114.
- Xue, C., Seetharam, AS., and Musharova, O. (2015). CRISPR interference and priming varies with individual spacer sequences. *Nucleic acids research*. 43, 10831-10847. doi: 10.1093/nar/gkv1259.
- Yang, H., Wang, H., Shivalila, C. S., Cheng, A. W., Shi, L., and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*, 154(6), 1370-1379. <https://doi.org/10.1016/j.cell.2013.08.022>.
- YANG, H., ZHANG, J., ZHANG, X., SHI, J., PAN, Y., ZHOU, R., LI, G., LI, Z., CAI, G. and WU, Z. (2018). CD163 knockout pigs are fully resistant to highly pathogenic porcine reproductive and respiratory syndrome virus. *Antiviral research*. <https://doi.org/10.1016/j.antiviral.2018.01.004>
- Yang, J., Zimmerly, S., Perlman, PS., and Lambowitz, AM. (1996). Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature*. 381, 332. doi:10.1038/381332a0.
- Yang, Y., Wang, L., Bell, P., McMenamin, D., He, Z., White, J., ... and Batshaw, M. L. (2016). A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn

- mice. *Nature biotechnology*, 34, 334. doi:10.1038/nbt.3469.
- Yin, H., Xue, W., and Chen, S. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* 32, 551-553. doi: 10.1038/nbt.2884.
- Yosef, I., Goren, MG., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic acids research*. gks216. doi: 10.1093/nar/gks216.
- Zaleski, P., Wojciechowski, M., and Piekarowicz, A. (2005). The role of Dam methylation in phase variation of *Haemophilus influenzae* genes involved in defence against phage infection. *Microbiology*. 151, 3361-3369. DOI:10.1099/mic.0.28184-0.
- Zetsche, B., Gootenberg, JS., Abudayyeh OO, et al., (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. 163, 759-771. doi: 10.1016/j.cell.2015.09.038.
- Zhang, F., Cong, L., and Lodato, S. (2011). Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* 29, 149-153. doi: 10.1038/nbt.1775.
- Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S., and Yang, S. H. (2015). Off-target effects in CRISPR/Cas9-mediated genome engineering. *Molecular Therapy-Nucleic Acids*. 4, 2162-2531. doi:10.1038/mtna.2015.37
- Zhang, X., Wang, J., Cheng, Q., Zheng, X., Zhao, G., and Wang, J. (2017). Multiplex gene regulation by CRISPR-ddCpf1. *Cell discovery*, 3, 17018. doi:10.1038/celldisc.2017.18
- Zhang, Y., Heidrich, N., and Ampattu, BJ. (2013). Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell*. 50, 488-503. doi: 10.1016/j.molcel.2013.05.001.
- Zhao, C., Zheng, X., Qu, W., Li, G., Li, X., Miao, Y. L., ... and Shao, Q. (2017). CRISPR-offinder: a CRISPR guide RNA design and off-target searching tool for user-defined protospacer adjacent motif. *International journal of biological sciences*, 13(12), 1470. doi: 10.7150/ijbs.21312.
- Zimmerly, S., Guo, H., and Eskest, R. (1995). A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell*. 83, 529-538. [https://doi.org/10.1016/0092-8674\(95\)90092-6](https://doi.org/10.1016/0092-8674(95)90092-6).
- Zuris, J. A., Thompson, D. B., Shu, Y., Guilinger, J. P., Bessen, J. L., Hu, J. H., ... and Liu, D. R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nature biotechnology*, 33, 73-80. doi:10.1038/nbt.3081.

Further reading:

The CRISPR/Cas System: Emerging Technology and Application

"reviews recent advances" (ProtoView)

Publisher: Caister Academic Press
Edited by: Muhammad Jamal

DOI: <https://doi.org/10.21775/9781910190630>

<https://www.caister.com/crispr>