BMB
Reports

BMB Rep. 2020; 53(7): 341-348
www.bmbreports.org

Invited Mini Review

Development of CRISPR/Cas9 system for targeted DNA modifications and recent improvements in modification efficiency and specificity

*Juhyun Shin & Jae-Wook Oh**
Department of Stem Cell and Regenerative Biotechnology, Konkuk University, Seoul 05029, Korea

The targeted nuclease clustered, regularly interspaced short palindromic repeats/CRISPR-associated proteins (CRISPR/Cas) system has recently emerged as a prominent gene manipulation method. Because of its ease in programming targeted DNA/protein binding through RNA in a vast range of organisms, this prokaryotic defense system is a versatile tool with many applications in the research field as well as high potential in agricultural and clinical improvements. This review will present a brief history that led to its discovery and adaptation. We also present some of its restrictions, and modifications that have been performed to overcome such restrictions, focusing specifically on the most common CRISPR/Cas9 mediated non-homologous end joint repair. [BMB Reports 2020; 53(7): 341-348]

INTRODUCTION

Discovered as an immune system against viral infection in domain bacteria and archaea, clustered regularly interspaced short palindromic repeats (CRISPR) system has quickly become a crucial tool in biological research. Long before it became the focus of debate because of its use to generate gene-edited babies (1), scientists recognized CRISPR system as an efficient, accurate and programmable nuclease system capable to induce double strand breaks (DSBs) in various organisms, therefore with a high potential as a versatile tool for scientific studies as well as a powerful tool for medicinal and agricultural improvement. Already, there are abundant data on the potential of CRISPR system for application in crop improvement (2) as an alternative to genetically modified organisms (GMO) (3). Moreover, clinical trials to cure cancer patients by using CRISPR edited T cells are ongoing in the United States (ClinicalTrials.gov registry number: NCT03399448) and China (registry number: NCT03545815).

*Corresponding author. Tel: +82-2-2049-6271; Fax: +82-2-455-1044; E-mail: ohjw@konkuk.ac.kr

https://doi.org/10.5483/BMBRep.2020.53.7.070

Received 31 March 2020

Keywords: Cas9, CRISPR, DSB, NHEJ, Targeted mutation

GENE EDITING PRIOR TO CRISPR/Cas9 SYSTEM

From restriction enzyme technique (Fig. 1A), the ability to manipulate genomic DNA in living cells had a pivotal role in the history of biological research. In the 1980s, directed mutation technique via homologous recombination revolutionized the field by allowing directed mutation in mammalian cells (4). Homologous recombination is performed by introducing trans-acting DNA material, usually containing a selection marker, by flanking homologous sequence matching target genomic DNA (Fig. 1B). This technology led to pivotal discoveries at that time. As an example, directed gene deletion by homologous recombination of mice stem cells allowed subsequent generation of transgenic mouse bearing deletion in genes of interest (5). However, this technique efficiency was characteristically low (6). Thus, further researches focused on overcoming those limitations, for example by using negative selection markers such as thymidine kinase or diphtheria toxin fragment A. However, efficiency increase remained modest [reviewed in (7)]. In the 1990s, Rouet, et al. discovered that *I-SceI*, a *Saccharomyces cerevisiae* derived endonuclease, can introduce a DSB in mammalian cells (8). This leaded to the discovery of homing endonucleases that can be used to induce lateral transfer of an intervening sequence via DSB (Fig. 1C) (9). One of the major limitations with these homing endonucleases was that their recognition sites are relatively longer than most restriction enzymes (14-40 bp), limiting the number of suitable targets, prompting researches aiming to diversify homing enzymes recognition sites (10-12).

In 1992, *Fok I*, another nuclease from *Flavobacterium okeanokoites* came to focus as its active domain and binding domain was identified (13). Interestingly, *Fok I* active domain requires homodimerization for activation, that is a pair of DNA-protein binding to flanking site to trigger its nuclease activity. Conversely, zinc finger motifs were known as short motifs in regulatory proteins that only need a three bases recognition site to bind the DNA (14). Thus, fusion of zinc finger DNA binding domain to *Fok I* nuclease domain (15, 16)

ISSN: 1976-670X (electronic edition)

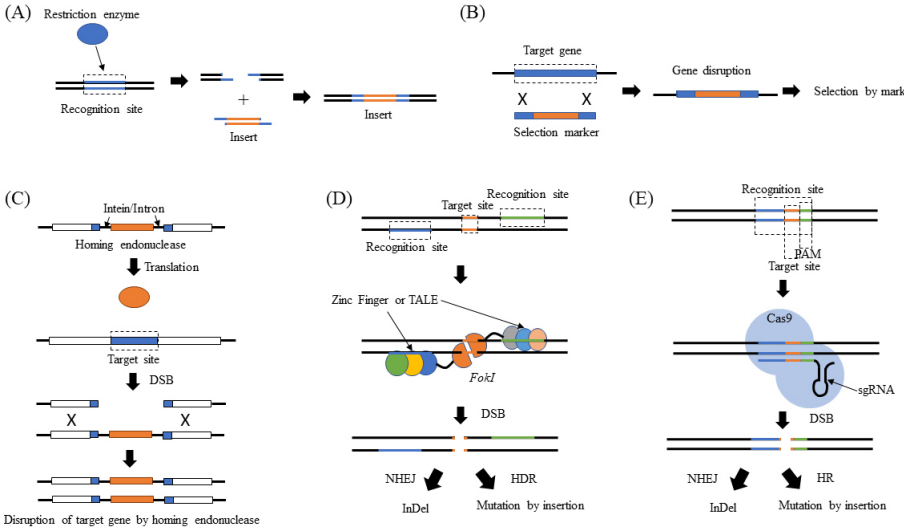

Fig. 1. Representative cartoon of the evolution of the directed mutation technique from enzyme restriction (A), homologous recombination (B), homing endonuclease (C), ZFN/TALEN (D) and the CRISPR/Cas9 system (E).

allowed development of chimeric enzymes, or zinc finger nucleases (ZFNs), that can induce targeted DSB as a pair (17). Gene manipulation of the ZFNs domain allowed targeted gene editing in a broad range of organisms, opening novel experimental and therapeutic possibilities (18). Later in 2009, transcription activator-like effector (TALE) proteins from the plant pathogenic bacteria *Xanthomonas* was shown to have two hypervariable amino acid residues that can recognize a single base pair (19), leading to the development of a chimeric protein with TALE binding domain and *Fok I* nuclease domain, or TALEN (Fig. 1D) (20). While the development of the ZNFs and TALENs allowed targeted gene manipulation in living cells, because of theirs mandatory cloning and protein modification steps needed to program target loci, the use of these techniques were still restricted. In contrast, the newly emerging CRISPR/Cas9 technology uses a short guide RNA to direct its endonuclease activity, that is more convenient to manipulate (Fig. 1E).

DISCOVERY AND ELUCIDATION OF THE CRISPR/CAS SYSTEM

CRISPR system is originally an RNA-mediated defense system in bacteria and archaea (21). While first discovered in *Escherichia* (*E.*) *coli* in the 1980s (22), the term CRISPR was later coined as its repeat motif was identified among other prokaryotes (23). CRISPR motif contains several nearly palindromic 30 bp repeats interspaced by ~36 bp non-repetitive spacers, with *Cas* genes nearby (24). In 2005, three independent reports showed that CRISPR various spacers are present in various prokaryotes strains, including among others, *Streptococcus*, *Sulfololus*, *Eschericia* and *Listesria* genus, and that those spacers are from mobile genetic elements such as viruses and phages, implying that CRISPR motifs are part of a major prokaryotic defense mechanism toward foreign genetic elements (25-27).

In 2007, Barrangou, et al. reported that CRISPR spacers provide prokaryotes resistance against corresponding phage by mechanisms involving neighbor *Cas* genes (28). Following this key experiment, researchers quickly started to elucidate the mechanism of CRISPR/Cas system. In a CRISPR/Cas system inserted in *E. coli*, it was shown that a complex of five Cas proteins is required for maturation of a 61 bp CRISPR RNA (crRNA) that comprise a spacer flanked by two repeat sequences (29). This led to an early hypothesis that crRNA may form a secondary structure (30). One of the first speculations was that these small RNAs act by similar mechanism to the well-studied small RNA interfering system, that is by RNA-RNA interaction between spacer and target RNA leading to foreign RNA degradation (25). However, modifying the intron sequence within the target sequence abolished the CRISPR/Cas defense system, thus demonstrating that the target of crRNA is not mRNA but DNA (31). In the same year, it was shown that the repeats that are adjacent to spacers are critical for the *S. thermophilus* defense mechanism (32, 33) and that Cas9, a *S. thermophilus* Cas protein, cleaved the foreign DNA at exactly the same position relative to these repeats that were later named Protospacer Adjacent Motif (PAM) (34). Various eukaryotes have specific PAM sequences that are recognized by specific Cas genes. Further investigations were focused on *S. pyogenes* Cas 9 (spCas9) based CRISPR/Cas9 system, as it requires a relatively short 5′-NGG PAM sequence for its activity.

Parallel to these studies, another focus of this field was to elucidate the mechanism underlying crRNA maturation, as many organisms with a working CRISPR/Cas9 system lacked essential Cas proteins thought to be essential for crRNA maturation in previous studies. Trans-activating CRISPR RNA (tracrRNA) were first identified as the third most abundant

class of transcripts in the *S. pyogenes*. As their name suggests, these small RNAs transcribed from locus adjacent to the CRISPR, and more importantly, were shown to be essential for crRNA maturation (35), and later on, for Cas9 nuclease activity. Thus, tracrRNAs were recognized as essential components of this system (36, 37).

EMERGENCE OF CRISPR/Cas9 AS A GENOME EDITING TOOL

In 2011, as all the crucial components of the CRISPR/Cas9 systems were identified, Siksnys' lab successfully reconstructed a working CRISPR/Cas9 system in *E. coli* from *S. thermophilus*, demonstrating for the first time that the CRISPR is transferable between organisms (38). Subsequently in 2012, two independent groups demonstrated the potential of the CRISPR/Cas9 system as a biological tool. Siksnys and his colleagues showed that *in vitro*, purified his-tagged Cas9 and custom designed spacers can introduce a DSB at a locus that is 3 bases away from PAM site (39). Conversely, Charpentier and her colleagues showed that the cut site is programmable, using a *S. pyogenes* Cas9 expressed in *E. coli* with *in vitro* transcribed crRNA and tracrRNA, and also a fused single-guide RNA (sgRNA), which is now widely used in gene editing. sgRNA is a crRNA and tracrRNA hybrid RNA that comprises a stem loops structure, repeat versus anti-repeat duplex, a *S. pyogenes* specific PAM site (5'-NGG-3') directly adjacent to 20 bp complementary to the target sequence (36). Both groups recognized the potential of their results, stating the potential use of this system as an RNA programmable DNA editing technique. Further studies showed that the CRISPR can be adapted *in vivo* for eukaryotes, notably human cells (40-43). Shortly after its potential as a genome editing tool was presented, the CRISPR/Cas9 system was widely adopted by the scientific community because of its ease in programming, yet high specificity to perform gene editing at target sites, with more than 15,000 papers published, and more than 4,500 in 2019 (Fig. 2).

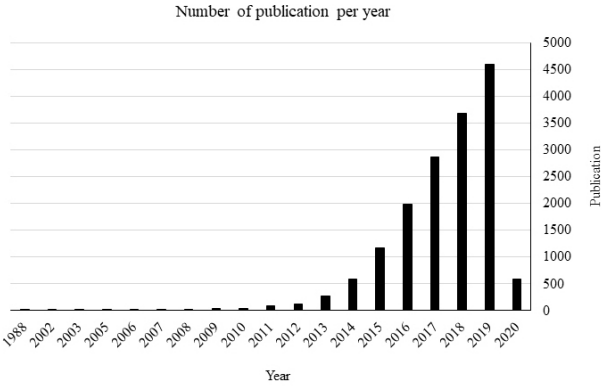

Fig. 2. Graph depicting the number of the CRISPR/Cas9 related publications yearly by MeSH (Medical Subject Headings) 1980 to March, 2020.

CRISPR/CAS9 OFF-TARGET EFFECT

While the CRISPR/Cas9 system is quickly becoming the tool of choice to perform targeted mutation, there are still limitations to address. One of these was raised and assessed early from the groups that developed the CRISPR technology (44-47). Off-target, that is nuclease activity at sites other than the programmed sites, can occur in the CRISPR/Cas9 system [the widely used CRISPR system, Class 2 type II from the *S. pyogenes* (48)] if the sgRNA binding allows mismatches. Cleavage occurs after the third nucleotide from the PAM within the corresponding target sequence (49). It was shown that sgRNA mismatch can be tolerated up to five base pairs, depending on the sgRNA and Cas9 amount used (45). Moreover, it was demonstrated by chromatin immunoprecipitation-sequencing (ChIP-seq) that binding sites of the deactivated Cas9 (dCas9)/sgRNA complex are less stringent after the fifth base adjacent to the PAM site, resulting in multiple off-target bindings (50, 51), preferably in the open chromatin region (52). However, while the dCas9/sgRNA allows non-specific binding, it was suggested that the guide RNA and the target PAM distal site sequence interaction is necessary for the Cas9 cleavage activity (53), thus decreasing the occurrence of off-targets *in vivo* compared to the occurrence of binding identified by the ChIP-seq. Indeed, relative to the large number of off-target dCas9 binding loci, only few or no insertion-deletions mutation (InDels) were observed at these off-target sites (49, 51).

IMPROVEMENT TO REDUCE OFF-TARGET

Researchers investigated various paths toward improving the Cas9 on-target efficiency. A straightforward method was to optimize Cas9 and gRNA amount and proportion (45), or directly deliver the Cas9 protein and sgRNA as ribonucleoprotein into cells (54). Additionally, while studies of the Cas9 variants from other organisms than the *S. pyogenes* have mostly contributed to broaden the scope of CRISPR technology application by diversifying available the PAM motifs (55), some of those variants may be also be used to optimize target specificity (56). While the consensus is that the probability to discover new subtypes more efficient than the known Cas9 is low, deeper understanding of Cas proteins variants and discovery of related proteins may contribute to technical enhancements (57). For example, Cas12a, formally known as Cpf1 (58), is annotated as a class 2 type V CRISPR system (59, 60). It requires only a single crRNA to introduce a staggered DSB, thus requires a simpler guide RNA and can be used to control insert orientation by its staggered DSB. CRISPR/Cpf1 has been used to generate targeted knock-out mice without any off-target effects observed, suggesting that it has potential as a DNA editing tool with efficiency that is at least comparable to

the Cas9 system (61). Some of the Cas9 variants even shows to have increased specificity of targeting site, among other improvements such as PAM diversification and increase of efficiency (Table 1). However, it should be noted that off-target should still be carefully considered, as a recent study using reporter activation assay to investigate editing efficiency shows that both CRISPR/Cas9 and CRISPR/Cpf1 tolerate off-target mismatch mutation, especially in PAM-distal region of investigated target (63).

Recently, new techniques using dCas9 variants fused to proteins are emerging as promising tools in gene editing. For example, base editing is a technique that uses dCas9 fused to a nucleobase deaminase enzyme or a DNA glycosylase able to convert a single base pair in targeted site, enabling precise point mutation (64). Prime editing, or search-and-replace genome editing is a technique that uses a dCas9 fused to a reverse transcriptase domain and a modified gRNA to insert a designed sequence within target site, therefore enabling precise sequence insertion without donor DNA (65). While they do not induce DSBs, these editing techniques are also prone to the same off-target issue as CRISPR/Cas9 system. Indeed, it was shown that base editing occurs at off-target sites in a frequency ranging from 0.07% to 100% in 38-58% genes in human cell (66). Interestingly, it was shown that improvement in base editing on-target efficiency can be achieved by optimizing the base editing domain rather than the dCas9 domain (67).

Finally, optimization can be achieved by sgRNA configuration. The five base pairs at proximity of the PAM region of the sgRNA are known as 'seed regions' more stringent in guiding the Cas9 complex to its target (51). While distal sequences from the PAM region are necessary to the Cas9 activity (53), those 'seed regions' context are crucial in determining the binding specificity. For example, U-rich seeds sequence increases specificity (51, 68) while high GCs content decreases Cas9 activity (68, 69). G is more favorable and C is less favorable as the base directly after the PAM. In contrast, at the fifth base from the PAM site, C is more preferred. From the 9th to 10th distal sequence from the PAM, A is favorable. At the 18th base from PAM, C is less favorable (51, 68, 70, 71). Those criteria can be used to design a target site for the gene of interest with minimum putative off-targets.

OFF-TARGET DETECTION

In order to use CRISPR/Cas9 system as a therapeutic tool, common agreement is that the risk of off-target should be assayed in a case specific manner. The impact of off-target mutation in patients will differ greatly based on the genes and tissues affected, as factors such as differential expression between tissues and pathological effect of genes differ greatly (72). Thus, to enable practical application, another focus was the development of tools that facilitate analysis of the off-targets using whole-genome sequencing (WGS) with improvement in cost and efficiency. To this end, several methods were developed including BLESS/BLISS which label and detect breaks *in situ* (73), Guide-Seq which incorporates oligo within DSBs as priming targets (74), Digenome-seq that perform WGS in nuclease digested sample to detect random modification (75), qDSB-Seq which compare DSBs 'spike' between two samples upon random DSBs (76) and recently DISCOVER-seq which track Cas9 binding site using Cas9-ChIP and WGS (77).

Table1. List of Cas9 variants developed to enhance specificity of sgRNA targeting

Name	Modification	Note
Cas9n or D10A Cas9 nickase	D10A	Developed by targeted mutation to induce rare DSB from nucleotide nickase (38), its paired D10A usage has reduced off-target and increased efficiency (62)
SpCas9-HF	N497A, R661A, Q695A, Q926A	Alteration of amino acid at gDNA interacting domain increase specificity (57) based on structural data from Cas9/gDNA complex crystallization (59)
espCas9(1.0)/(1.1)	K810A/K848A, K1003A, R1060A	Modification of amino acid at interacting with noncomplementary DNA strand based on structural data from Cas9/gDNA complex crystallization (61)
HypaCas9	N692A, M694A, Q695A, H698A	Alteration in REC3 domain, identified by single-molecule Förster resonance energy transfer experiments (68)
HifiCas9	R691A	Identified by unbiased bacterial screening method (69)
evoCas9	M495V, Y515N, K526E, R661Q	Yeast based screening of random mutation in the REC3 domain (70)
SniperCas9	F5395, M763I, K890N	Directed evolution in a *E. coli* system to screen for accurate and efficient nuclease (71)
xCas9-3.6	E108G, S217A, A262T, S409I, E480K, E543D, M694I, E1219V	Phage assisted evolution to screen for Cas9 variants to diversify PAM sites. Those variants are also more specific to target sites (72)
xCas9-3.7	A262T, R324L, S409I, E480K, E543D, M694I, E1219V	

ON-TARGET GENE SILENCING BY CRISPR/Cas9

Another putative side effect of the targeted CRISPR/Cas9 system is on-target mis-regulation. The underlying mechanism of gene mutation by CRISPR/Cas9 systems is that Cas9 induces DSB in the genome that triggers repair pathways via the non-homologous joint end repair (NHEJ) or the less frequent homology directed repair (HDR) that occurs if a homologous template is nearby (78). In eukaryote, DSBs occur relatively frequently because of reactive oxygen species, radiation, replication error or mechanical stress. Thus, proteins that are involved NHEJ are intrinsically active (79). Repair by NHEJ usually results in imperfect repairs with insertion or deletion mutations (InDels), leading to a frameshift mutation that consequently results to a premature termination codon (PTC) within coding region. PTC triggers cell inherent nonsense-mediated mRNA decay (NMD) mechanism leading to complete knock out of the targeted gene by mRNA degradation within seconds (80, 81). On-target CRISPR/Cas9 mediated gene silencing is usually achieved by this mechanism.

ON-TARGET MISREGULATION AND IMPROVEMENT

However, a recent investigation showed that ~50% of the cell lines from previous studies did not result in targeted gene knock out, but rather caused the production of truncated functional proteins. To reduce on-target mis-regulation, the authors recommended selecting target sites avoiding the internal ribosomal re-entry site, as InDels in those sites may result in the production of pseudo-mRNA. Also, exon splicing enhancers site should be avoided as target site as their deletion may result in exon skipping, thus generating truncated proteins rather than knock out (82). While other studies applied this exon skipping capability to introduce alteration in the targeted genes (83-86), the consensus is that in addition to the off-target mutations, these on-target mis-regulations should be carefully evaluated before application (84, 86).

CONCLUSION

For now, CRISPR/Cas9 system is known as the most convenient method to program target sites for mutation among developed techniques. Thus, CRISPR/Cas9 system is quickly becoming a prominent tool for basic research as well as for clinical and agricultural purposes. In this review, we discussed a few of the many studies that led to its development. Its basic principle is that it induce a targeted DBS in the genome that can go through two inherent mechanisms, NHEJ that ligate the break without a homologous template and HDR that use a homologous template, therefore that is less error-prone but has lower efficiency compared to NHEJ. Thus, NHEJ remains the most commonly used pathway despite its putative on/off-target side effect. Recent improvements have been initiated to increase the specificity of the Cas9 targeted DSB as well as to develop techniques to detect off-target at large scale, crucial to evaluating its safety for clinical and agricultural applications. While off-target mutation can be detected using large-scale analyses, on-target mis-regulation can only be assessed after mutation in a case specific manner. This shows that the CRISPR/Cas9 possible side-effects should be carefully assayed before application, and there is room for improving this highly effective targeted mutation technique.

ACKNOWLEDGEMENTS

This study was supported by grants awarded to the JWO (NRF-2017H1D3A1A01052995 and NRF-2016R1D1A1B03935382) and the JS (NRF-S201806S00067) by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology. We apologize to colleagues whose work could not be cited because of space limitations.

CONFLICTS OF INTEREST

The authors have no conflicting interests.

REFERENCES

1. Cyranoski D (2019) The CRISPR-baby scandal: what's next for human gene-editing. Nature 566, 440-442
2. El-Mounadi K, Morales-Floriano ML and Garcia-Ruiz H (2020) Principles, Applications, and Biosafety of Plant Genome Editing Using CRISPR-Cas9. Front Plant Sci 11, 56
3. Lassoued R, Macall DM, Hesseln H, Phillips PWB and Smyth SJ (2019) Benefits of genome-edited crops: expert opinion. Transgenic Res 28, 247-256
4. Smithies O, Gregg RG, Boggs SS, Koralewski MA and Kucherlapati RS (1985) Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature 317, 230-234
5. Bollag RJ, Waldman AS and Liskay RM (1989) Homologous recombination in mammalian cells. Annu Rev Genet 23, 199-225
6. Vasquez KM, Marburger K, Intody Z and Wilson JH (2001) Manipulating the mammalian genome by homologous recombination. Proc Natl Acad Sci U S A 98, 8403-8410
7. Bouabe H and Okkenhaug K (2013) Gene targeting in mice: a review. Methods Mol Biol 1064, 315-336
8. Rouet P, Smih F and Jasin M (1994) Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A 91, 6064-6068
9. Chevalier BS and Stoddard BL (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucleic Acids Res 29, 3757-3774
10. Rosen LE, Morrison HA, Masri S et al (2006) Homing endonuclease I-CreI derivatives with novel DNA target specificities. Nucleic Acids Res 34, 4791-4800

11. Seligman LM, Chisholm KM, Chevalier BS et al (2002) Mutations altering the cleavage specificity of a homing endonuclease. Nucleic Acids Res 30, 3870-3879
12. Sussman D, Chadsey M, Fauce S et al (2004) Isolation and characterization of new homing endonuclease specificities at individual target site positions. J Mol Biol 342, 31-41
13. Li L, Wu LP and Chandrasegaran S (1992) Functional domains in Fok I restriction endonuclease. Proc Natl Acad Sci U S A 89, 4275-4279
14. Choo Y and Klug A (1994) Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. Proc Natl Acad Sci U S A 91, 11168-11172
15. Kim YG, Cha J and Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93, 1156-1160
16. Kim YG, Shi Y, Berg JM and Chandrasegaran S (1997) Site-specific cleavage of DNA-RNA hybrids by zinc finger/FokI cleavage domain fusions. Gene 203, 43-49
17. Bibikova M, Beumer K, Trautman JK and Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. Science 300, 764
18. 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
19. Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326, 1509-1512
20. Christian M, Cermak T, Doyle EL et al (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186, 757-761
21. Mojica FJ, Diez-Villasenor C, Soria E and 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
22. 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
23. Jansen R, Embden JD, Gaastra W and Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol 43, 1565-1575
24. Mojica FJ, Juez G and Rodriguez-Valera F (1993) Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. Mol Microbiol 9, 613-621
25. 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
26. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J and Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 60, 174-182
27. 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
28. Barrangou R, Fremaux C, Deveau H et al (2007) CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709-1712
29. Brouns SJ, Jore MM, Lundgren M et al (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960-964
30. Sorek R, Kunin V and Hugenholtz P (2008) CRISPR–a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat Rev Microbiol 6, 181-186
31. Marraffini LA and Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 322, 1843-1845
32. Deveau H, Barrangou R, Garneau JE et al (2008) Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J Bacteriol 190, 1390-1400
33. Horvath P, Romero DA, Coute-Monvoisin AC et al (2008) Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. J Bacteriol 190, 1401-1412
34. Garneau JE, Dupuis ME, Villion M et al (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67-71
35. Deltcheva E, Chylinski K, Sharma CM et al (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471, 602-607
36. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA and Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821
37. Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P and Siksnys V (2013) crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. RNA Biol 10, 841-851
38. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P and Siksnys V (2011) The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Res 39, 9275-9282
39. Gasiunas G, Barrangou R, Horvath P and Siksnys V (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci U S A 109, E2579-2586
40. Cong L, Ran FA, Cox D et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823
41. Jinek M, East A, Cheng A, Lin S, Ma E and Doudna J (2013) RNA-programmed genome editing in human cells. Elife 2, e00471
42. Mali P, Yang L, Esvelt KM et al (2013) RNA-guided human genome engineering via Cas9. Science 339, 823-826
43. Cho SW, Kim S, Kim JM and Kim JS (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31, 230-232
44. Cho SW, Kim S, Kim Y et al (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 24, 132-141
45. Fu Y, Foden JA, Khayter C et al (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31, 822-826
46. Hsu PD, Scott DA, Weinstein JA et al (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nat

Biotechnol 31, 827-832
47. Mali P, Aach J, Stranges PB et al (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31, 833-838
48. Koonin EV, Makarova KS and Zhang F (2017) Diversity, classification and evolution of CRISPR-Cas systems. Curr Opin Microbiol 37, 67-78
49. Anders C, Niewoehner O, Duerst A and Jinek M (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature 513, 569-573
50. 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. Nat Biotechnol 32, 677-683
51. Wu X, Scott DA, Kriz AJ et al (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol 32, 670-676
52. Singh R, Kuscu C, Quinlan A, Qi Y and Adli M (2015) Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. Nucleic Acids Res 43, e118
53. Cencic R, Miura H, Malina A et al (2014) Protospacer adjacent motif (PAM)-distal sequences engage CRISPR Cas9 DNA target cleavage. PLoS One 9, e109213
54. Kim S, Kim D, Cho SW, Kim J and Kim JS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res 24, 1012-1019
55. Pickar-Oliver A and Gersbach CA (2019) The next generation of CRISPR-Cas technologies and applications. Nat Rev Mol Cell Biol 20, 490-507
56. Cebrian-Serrano A and Davies B (2017) CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. Mamm Genome 28, 247-261
57. Wright AV, Nunez JK and Doudna JA (2016) Biology and applications of CRISPR systems: Harnessing Nature's Toolbox for Genome Engineering. Cell 164, 29-44
58. Makarova KS, Wolf YI and Koonin EV (2018) Classification and nomenclature of CRISPR-Cas systems: where from here? CRISPR J 1, 325-336
59. 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
60. Paul B and Montoya G (2020) CRISPR-Cas12a: Functional overview and applications. Biomed J 43, 8-17
61. Kim Y, Cheong SA, Lee JG et al (2016) Generation of knockout mice by Cpf1-mediated gene targeting. Nat Biotechnol 34, 808-810
62. Makarova KS, Wolf YI, Alkhnbashi OS et al (2015) An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol 13, 722-736
63. Wang Y, Wang M, Zheng T et al (2020) Specificity profiling of CRISPR system reveals greatly enhanced off-target gene editing. Sci Rep 10, 1-8
64. Rees HA and Liu DR (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. Nat Rev Genet 19, 770-788
65. Anzalone AV, Randolph PB, Davis JR et al (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 576, 149-157
66. Grunewald J, Zhou R, Garcia SP et al (2019) Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. Nature 569, 433-437
67. Grunewald J, Zhou R, Iyer S et al (2019) CRISPR DNA base editors with reduced RNA off-target and self-editing activities. Nat Biotechnol 37, 1041-1048
68. Wang T, Wei JJ, Sabatini DM and Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80-84
69. Doench JG, Hartenian E, Graham DB et al (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32, 1262-1267
70. Gagnon JA, Valen E, Thyme SB et al (2014) Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS One 9, e98186
71. Zhang XH, Tee LY, Wang XG, Huang QS and Yang SH (2015) Off-target effects in CRISPR/Cas9-mediated genome engineering. Mol Ther Nucleic Acids 4, e264
72. Carroll D (2019) Collateral damage: benchmarking off-target effects in genome editing. Genome Biol 20, 114
73. Mirzazadeh R, Kallas T, Bienko M and Crosetto N (2018) Genome-wide profiling of DNA double-strand breaks by the BLESS and BLISS methods. Methods Mol Biol 1672, 167-194
74. Tsai SQ, Zheng Z, Nguyen NT et al (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol 33, 187-197
75. Kim D, Bae S, Park J et al (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat Methods 12, 237-243, 231 p following 243
76. Zhu Y, Biernacka A, Pardo B et al (2019) qDSB-Seq is a general method for genome-wide quantification of DNA double-strand breaks using sequencing. Nat Commun 10, 1-11
77. Wienert B, Wyman SK, Yeh CD, Conklin BR and Corn JE (2020) CRISPR off-target detection with DISCOVER-seq. Nat Protoc 15, 1775-1799
78. Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79, 181-211
79. Nambiar TS, Billon P, Diedenhofen G et al (2019) Stimulation of CRISPR-mediated homology-directed repair by an engineered RAD18 variant. Nat Commun 10, 3395
80. Brogna S and Wen J (2009) Nonsense-mediated mRNA decay (NMD) mechanisms. Nat Struct Mol Biol 16, 107-113
81. Popp MW and Maquat LE (2016) Leveraging rules of nonsense-mediated mRNA decay for genome engineering and personalized medicine. Cell 165, 1319-1322
82. Tuladhar R, Yeu Y, Tyler Piazza J et al (2019) CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. Nat Commun 10, 4056
83. Winter J, Luu A, Gapinske M et al (2019) Targeted exon skipping with AAV-mediated split adenine base editors. Cell Discov 5, 1-12
84. Mou H, Smith JL, Peng L et al (2017) CRISPR/Cas9-mediated genome editing induces exon skipping by alter-

native splicing or exon deletion. Genome Biol 18, 1-8
85. Chen D, Tang JX, Li B, Hou L, Wang X and Kang L (2018) CRISPR/Cas9-mediated genome editing induces exon skipping by complete or stochastic altering splicing in the migratory locust. BMC Biotechnol 18, 1-9
86. Sui T, Song Y, Liu Z et al (2018) CRISPR-induced exon skipping is dependent on premature termination codon mutations. Genome Biol 19, 164