Living with CRISPR:

The origins of CRISPR and how does it function

Junhui Liu

Xi'an Jiaotong University

Abstract

In the last few decades, CRISPR is emerging as an indispensable tool in biological research. It was accidentally discovered from *Haloferax mediterranei*, then scholars scrutinized it for its specific biological role as a bacterial immune system against invading viruses, as well as for its more specific molecular working mechanisms. Consequently, the powerful programmability of the Cas9 enzyme in the CRISPR system was discovered and applied in a wide, efficient, and innovative way. CRISPR is now revolutionizing diverse fields such as medical research, biotechnology, and agriculture. CRISPR Cas9 is no longer just a gene-editing tool, but has many innovative references: catalytically impaired inactive Cas9 applications, including gene regulation, epigenetic editing, chromatin engineering, and imaging, now exceed other gene editing capabilities. Here, I will present a brief history exploring the discovery of CRISPR, as well as some clever and innovative applications based on CRISPR/Cas9 technology.

key words: CRISPR, Cas9

Living with CRISPR:

The origins of CRISPR and how does it function

Nowadays, cutting-edge science has become more accessible to the public, for example, an awareness that some diseases cannot be eradicated but genetic therapeutic is commonly widespread. Gene, the exquisite double-helix structure in almost every cell in your body, is regarded to dominate everything of you. Apart from some inevitable genetic diseases and some macroscopic features distinguishing you from others, even how much air you can take in for every breathe and why you hate carrots can be explained by your unique gene. However, sometimes your gene might get into trouble. Substitutions, deletions, or additions at certain sites may bring your body to an awkward situation. On the good side, you may luckily get the opportunity to gain immunity to certain kinds of viruses, like AIDS. The probability for this good thing to happen is extremely small. More commonly, diseases like cancers may show up and consequently suck the life out of the patients.

Correspondingly, scientists worldwide have been investigating great efforts on gene editing, for its potential to cure various diseases. They manage to find magical tools to change several bases on certain situations, expecting to reconstruct the cells or the body to a more healthy one.

While in the past decade, as a revolutionizing technology, the CRISPR, discovered as a swift and precise method, has astonished more than the biology field, which shows underlying benefits in medical research and agriculture, as well. However, as a newborn and powerful methodology, many people even some biologists remain superficially understanding it no more than general mechanisms and simple applications. When it comes to the overall development of CRISPR and essential molecular mechanisms for it to work, the majority might admit that they are uncertain about it. For instance, there are plenty of other abilities of

CRISPR you may not know, including gene regulation, epigenetic editing, chromatin engineering, and imaging.

Here, to reverse the great revolution to a more approachable one, I will start with an introduction of CRISPR, how it was discovered and what efforts had been made to achieve some remarkable milestones. Then some vital mechanisms will be explained, laying the basis for the next part, some applications of the CRISPR.

Review of Literature

History of CRISPR

It's so common in the science field that a legendary success starts small, and CRISPR is one of those as well. The near 30 years story of CRISPR originates from an abnormal microbial repeat sequence, then more and more bricks have been added to the castle of CRISPR.

The discovery of CRISPR sequence. When researching on the way restriction enzymes cut the genome of *Haloferax mediterranei*, under different salt concentrations, Mojica found a curious structure—multiple copies of a near-perfect, roughly palindromic, repeated sequence of 30 bases, separated by spacers of roughly 36 bases—that did not resemble any family of repeats known in microbes (Mojica et al., 1993).

Then more and more such sequences were discovered by Mojica, which also attracted Mojica a lot. By 2000, Mojica had found CRISPR loci in 20 different microbes, and in 2002, he renamed the sequence to the clustered regularly interspaced palindromic repeats(CRISPR).

The basic function of CRISPR——an Adaptive Immune System. However, the function of CRISPR remained mysterious. Meanwhile, the hypotheses abounded, including gene regulation, replicon partitioning, DNA repair, and other roles.

In 2003, the interest in CRISPR was drawn to the spacers. Mojica meticulously extracted each spacer and utilized BLAST to analyze them. Benefiting from the expanding

DNA database, he surprisingly found some spacers are similar to the present viruses or conjugative plasmids reaching an assumption that CRISPR loci encode the instructions for an adaptive immune system protecting microbes against specific infections (Mojica et al., 2005).

After learning about CRISPR from Mojica in late 2002, Phillippe Horvath began to use it to genotype the strains of his lactic-acid bacteria. By late 2004, he noticed there is a correlation between spacers and phage resistance. Then in 2005, working with Rodolphe Barrangou and Sylvain Moineau, they tested the hypothesis of Mojica that CRISPR was an adaptive immune system. They used two kinds of bacteriophages to infect phage-sensitive *S.thermophilius* strain. By analyzing the resistant strains, unlike the classical resistance mutations, the survived strains had acquired phage-derived sequences at their CRISPR loci. (Barrangou et al., 2007). Moreover, the insertion of multiple spacers correlated with increased resistance. They had seen acquired immunity in action.

Meanwhile, they also studied the role of two *Cas* genes, several genes adjacent to CRISPR loci, *Cas7* and *Cas9*, concluding that Cas7 was involved in generating new spacers and repeats, while Cas9 presumably cut nucleic acids (Bolotin et al., 2005; Makarova et al., 2006)—was necessary for phage resistance.

Further improvement of CRISPR—— Programming and orientation of CRISPR. Out of curiosity, John van der Oost and his colleagues inserted an *E.coli* CRISPR system into another *E.coli* lacking its endogenous system. Thus, they created an opportunity to further work on the complex of five Cas proteins, termed Cascade. Then by knocking out each component individually, they showed that Cascade is required for cleaving a long precursor RNA, transcribed from the CRISPR locus, into 61-nucleotide-long CRISPR RNAs(crRNAs). They cloned and sequenced a set of crRNAs from the Cascade complex and found that all started with the last eight bases of the repeat sequence, followed by the complete spacer and the beginning of the next repeat region, supporting earlier suggestions that the palindromic nature of the repeats would lead to secondary structure formation in the crRNA (Sorek et al., 2008).

To prove that the crRNA sequences are responsible for CRISPR-based resistance, they set out to design the first artificial CRISPR arrays, programming them to target several genes in lambda(λ) phage. As they predicted, the new CRISPR sequence showed resistance to phage λ. The first experiment to design and synthesis a specific CRISPR succeeded.

However, one thing remains confusing, to decide the targeted situation of the Cascade complex, the authors had designed two versions of the CRISPR array—one in the anti-sense direction (complementary to both the mRNA and coding strand), one in the sense direction (same as the coding strand). The fact that the sense version worked strongly suggested that the target may not be mRNA, and the hypotheses that CRISPR targets DNA.

Later, Luciano Marraffini and Erik Sontheimer experimented to prove the hypothesis right. They turned to molecular biology for help. Cleverly, they modified the *Nes* gene in the plasmid targeted by the CRISPR system—inserting a self-splicing intron in the middle of its sequence. If CRISPR targeted mRNA, the change would not affect interference because the intronic sequence would be spliced out. If CRISPR targeted DNA, the insertion would

abolish interference because the spacer would no longer match. The results were clear: the target of CRISPR was DNA (Marraffini & Sontheimer, 2008).

Cas9 Is Guided by crRNAs and Creates Double-Stranded and discovery of tracrRNA——molecular mechanisms to explain how CRISPR works. In 2007, the fact that CRISPR is an adaptive immune system has been popularly accepted. But Sylvain Moineau, working with Daniso, still spent efforts to understand the specific mechanism by which CRISPR cleaves DNA.

The crucial problem, unexpectedly, is that CRISPR was normally so swift and efficient that they could hardly observe how the Cascade complex cleaves DNA. Fortunately, they accidentally found CRISPR conferred only partial protection against plasmid transformation by electroporation. In one such inefficient strain, they could see linearized plasmids persisting inside the cells. Somehow, the process of plasmid interference had been slowed down enough to observe the direct products of CRISPR's action (Garneau et al., 2010).

This strain allowed them to dissect the process of cutting, showing the cutting of the plasmid depended on the Cas9 nuclease. When they sequenced the linearized plasmids, they found a single precise blunt-end cleavage event 3 nucleotides upstream of the proto-spacer adjacent motif (PAM) sequence. After expanding their analysis, they showed that viral DNA is also cut in precisely the same position relative to the PAM sequence. Moreover, the number of distinct spacers matching a target corresponded to the number of cuts observed.

Their results showed definitively that Cas9's nuclease activity cut DNA at precise positions encoded by the specific sequence of the crRNAs.

TracrRNA is a small RNA that would come to be called *trans*-activating CRISPR RNA. Working on several regulatory RNAs in microbes—including one near the CRISPR locus of *Streptococcus pyogenes*, Emmanuelle Charpentier and Jörg Vogel. The novel small

RNA that was transcribed from a sequence immediately adjacent to the CRISPR locus (in the region that had caught Charpentier's attention) and had 25 bases of near-perfect complementary to the CRISPR repeats. The complementarity suggested that this tracrRNA and the precursor of the crRNAs hybridized together and were processed into mature products by RNaseIII cleavage. Genetic deletion experiments confirmed this notion, showing that tracrRNA was essential for processing crRNAs and thus for CRISPR function (Deltcheva et al., 2011). Subsequent biochemical studies showed that tracrRNA was not only involved in processing crRNA but was also essential for the Cas9 nuclease complex to cleave DNA (Gasiunas et al., 2012; Jinek et al., 2012). This process is a marvelous milestone, for all of the essential factors for CRISPR to serve humanity have been collected.

Reconstituting CRISPR in other environments——in a distant organism, in vitro, and in mammalian cells.

Reconstituting CRISPR in a distant organism. In the 1980s, Virginijus Siksnys was originally researching the then-hot field of restriction enzymes. Just as he was bored with restriction enzymes, he learned about the CRISPR system and became fascinated by it.

As a chemist, he felt that he would only understand CRISPR if he could reconstitute it in vitro. After collecting all of the necessary components of the CRISPR system. He and his team set out to see whether the CRISPR system from *S.thermophilus* could be reconstituted in fully functional form in *E.coli*. To their delight, they found that transferring the entire CRISPR locus was sufficient to cause targeted interference against both plasmid and bacteriophage DNA (Sapranauskas et al., 2011). Using their heterologous system, they also proved that Cas9 is the only protein required for interference and that its RuvC- and HNH-nuclease domains (Bolotin et al., 2005; Makarova et al., 2006) are each essential for interference.

Till then, the field had reached a critical milestone: the necessary and sufficient components of the CRISPR-Cas9 interference system—the Cas9 nuclease, crRNA, and tracrRNA—were now collected. The system had been completely constructed based on elegant bioinformatics, genetics, and molecular biology. It was now time to turn to precise biochemical experiments to try to confirm and extend the results in a test tube.

Reconstituting CRISPR in vitro. Before applying CRISPR in more complex and advanced cells, a vitro test is necessarily required. After successfully using the heterologous expression system in *E.coli*, Siksnys and his colleagues decided to conduct further experiments, studying the activity of CRISPR in a test tube (Gasiunas et al., 2012). They showed that the complex could cleave a DNA target in vitro, creating a double-stranded break precisely 3 nucleotides from the PAM sequence—matching the in vivo observations of Moineau and colleagues.

Most dramatically, they demonstrated that they could reprogram Cas9 with custom-designed spacers in the CRISPR array to cut a target site of their choosing in vitro. By mutating the catalytic residues of the HNH- and RuvC-nuclease domains, they also proved that the former cleaves the strand complementary to the crRNA while the latter cleaves the opposite strand. And, they showed that the crRNA could be trimmed down to just 20 nucleotides and still achieve efficient cleavage. Finally, Siksnys showed that the system could also be reconstituted in a second way—by combining purified His-tagged Cas9, in-vitro-transcribed tracrRNA and crRNA, and RNase III—and that both RNAs were essential for Cas9 to cut DNA (Gasiunas et al., 2012).

Around the same time, Charpentier was focusing on the biochemical characterization of CRISPR in Vienna, while Jeniffer Doudna had been a prestigious structural biologist and RNA expert at the University of California, Berkeley. The two scientists decided to join forces. They used recombinant Cas9 (from *S. pyogenes* expressed in *E. coli*) and crRNA and

tracrRNA that had been transcribed in vitro (Jinek et al., 2012). Like Siksnys, they showed that Cas9 could cut purified DNA in vitro, that it could be programmed with custom-designed crRNAs, that the two nuclease domains cut opposite strands, and that both crRNA and tracrRNA were required for Cas9 to function. In addition, they showed that the two RNAs could function in vitro when fused into a single-guide RNA (sgRNA). The concept of sgRNAs would become widely used in genome editing, after modifications by others to make it work efficiently in vivo. A few years later, Doudna would call the world's attention to the important societal issues raised by the prospect of editing the human germline.

Reconstituting CRISPR in mammalian cells. As an expert in neurobiology and molecular biology, Feng Zhang at that time dedicated himself to searching for a general way to program transcription factors to serve his field. After TALEs (TranscriptionActivator–Like Effectors) were deciphered, the Zhang team could find a way to activate, repress, or edit genes precisely. Still, he remained on the lookout for a better approach.

Then in 2011, he heard about CRISPR and instantly became captivated, setting out to create a version of S. thermophilus Cas9 for use in human cells. With one-years' optimization to explore how to increase the proportion of Cas9 that went to the nucleus and the best isoform of tracrRNA for mammalian cells, by mid-2012, he had a robust three-component system consisting of Cas9 from either S. pyogenes or S. thermophilus, tracrRNA, and a CRISPR array. Targeting 16 sites in the human and mouse genomes, he showed that it was possible to mutate genes with high efficiency and accuracy—causing deletions via non-homologous end-joining and inserting new sequences via homologous recombination with a repair template. he also tested a two-component system with the short sgRNA fusion described in their in vitro study. The fusion turned out to work poorly in vivo, cutting only a minority of loci with low efficiency, but he found that a full-length fusion that restored a critical 30 hairpin solved the problem (Cong et al., 2013).

Different CRISPR systems. As a consequence of environment evolution, the CRISPR systems have evolved into different types, with respective structures of CRISPR-associated (Cas) genes that are typically adjacent to the CRISPR array (Haft et al., 2005; Makarova et al., 2011). As more and more CRISPR systems have been discovered, the CRISPR library is expanding. Also, the classification system is still under optimization.

Universally, there were two general classes of CRISPR systems, each containing multiple CRISPR types. Class 1 contains type I and type III CRISPR systems, which are commonly found in Archaea. While class 2 contains the type II, IV, V, and VI CRISPR systems. The type II CRISPR-Cas9 system from *Streptococcus pyogenes* is the most widely used one for its simple NGG PAM sequence requirements. Researchers are still actively exploring other CRISPR systems to identify Cas9-like effector proteins with differences in their sizes, PAM requirements, and substrate preferences.

Re-engineering CRISPR-Cas9 tools. The Cas9 variants found in nature are usually large proteins, limiting their packaging and delivery into different cell types. For example, the widely used SpCas9 protein is 1,366 aa, making the delivery very challenging. So apart from keeping on exploring appropriate natural CRISPR systems, more and more molecular biologists are transferring their attention to re-engineering the already well-characterized Cas9 proteins. The main aspects can be concluded as to reduce the size of Cas9 nucleases, increase the fidelity and expand the targeting scope of Cas9 variants.

Although reducing the size of existing Cas9 may be challenging, the Cas9 PAM requirements and targeting specificity had to be optimized by several groups. In one such study, researchers used an unbiased selection strategy to identify variants of SpCas9 and SaCas9 with more relaxed PAM sequence requirements (Kleinstiver, Prew, Tsai, Nguyen, et al., 2015; Kleinstiver, Prew, Tsai, Topkar, et al., 2015).

In addition to expanding the targeting scope of CRISPR tools, researchers are actively developing novel ways to increase the targeting specificity of the CRISPR-Cas9 system. Understanding the extent of off-target effects of CRISPR-Cas9 targeting has been one major goal. Given that CRISPR systems have evolved as a defense system against viruses that tend to frequently mutate, a slightly less specific CRISPR system would be advantageous to bacteria. Indeed, the early efforts to understand CRISPR targeting specificity highlighted this fact and demonstrated that the system may potentially have off-target effects (Hsu et al., 2013; Mali et al., 2013).

The easiest approach to increasing the targeting specificity is changing the delivery method of the Cas9-sgRNA complex. In contrast to the plasmid-based method, direct delivery of Cas9-sgRNA as a ribonucleotide protein (RNP) complex results in more transient Cas9 activity and hence less off-target effects (Kim et al., 2014). Additionally, tandem targeting a locus with two separate sgRNAs, using small molecule chemicals, optical light, and ligand-dependent allosteric regulation to control activities of the Cas9/sgRNA complex to improve targeting specificity and modifying the sgRNA scaffold to increase the targeting specificity are three other ways under research.

Application of CRISPR

So far, the review has focused on the history and some basic mechanisms of CRISPR targeting and some of the recent approaches that have been utilized to monitor or improve the targeting specificity of CRISPR-Cas9. Due to its robustness and flexibility, CRISPR is becoming a versatile tool with applications that are transforming not only genome-editing studies but also many other genome and chromatin manipulation efforts.

CRISPR-Cas9 as a tool for gene therapy. Considered a promising solution to genetic diseases or cancers, gene therapy has gained quite a lot of hype in recent years. But there are still many people who believe that gene therapy is a huge scam. We have to admit that gene therapy raises some ethical issues as well as the lack of faster and more convenient means. But there are two sides to everything, and we cannot ignore the large benefits of gene therapy, for example, Andreas and his colleagues used gene therapy in 2019 to find an idea for treating exudative age-related macular degeneration (AMD) or cancers (Holmgaard et al., 2019).

They transferred two plasmid vectors into mouse retinal cells, one plasmid expressing *Streptococcus pyogenes* (Sp) Cas9 protein and the other plasmid expressing single guide RNAs (sgRNAs). After successfully expressed, they will be used as a CRISPR system to ablate the vascular endothelial growth factor A (*VEGFA*) gene in the mouse retina. Exudative AMD is mainly abnormal neovascularization in the subretinal growth of the macula, and *VEGFA* is a key regulator of angiogenesis, and notably, *VEGFA* is also crucial to regulate the growth process of solid tumors. These findings may help develop a new way to treat various diseases.

CRISPR-mediated gene expression regulation. Apart from targeting the CRISPR system to the DNA locus and editing the gene, dead Cas9 protein (dCas9) in the CRISPR system can also be a targeting protein to the specific locus. dCas9 strongly binds to the DNA target sequence and this tight binding interferes with the activity of other DNA binding proteins such as endogenous transcription factors and RNA Polymerase II. Consequently, with the dCas9 based complex, the gene expression can be enhanced (Thakore et al., 2015) or repressed (Cheng et al., 2013).

CRISPR-mediated epigenome editing. Epigenome modification means some heritable gene expression changes not attributed to changes in DNA sequence information, like post-translation modification and other chromatin features associated with regulatory elements in the genome. With the advantage of programmability, dCas9 can be modified to combine some epigenetic modification protein, thus, the function of epigenome can be further studied.

Harnessing the above principle, Marvin E. Tanenbaum used a nuclease-deficient Cas9 (dCas9)–histone demethylase fusion to functionally characterize previously described and new enhancer elements for their roles in the embryonic stem cell state (Tanenbaum et al., 2014). They created an expression system in mouse embryonic stem cells (mESCs), versions of Neisseria meningitidis (Nm) dCas9 are produced to target a certain locus with the help of sgRNAs, which are delivered into the cell by a viral system in advance, then a histone demethylase (LSD1) is combined to the dCas9 protein mouse embryonic stem cells (mESCs). In comparison with other epigenome editing methods, like non-effector BirA affinity tag (BAT) or a KRAB repressor, they could functionally characterize previously described and new enhancer elements for their roles in the embryonic stem cell state.

CRISPR-mediated live-cell chromatin imaging. Imaging chromatin dynamics is crucial to understanding genome organization and its role in transcriptional regulation. Using the sgRNA-guided CRISPR/dCas9 system to image chromatin in live cells can solve a tricky problem that through other methods, the cells observed are usually dead (Qin et al., 2017).

Qin designed single-guide RNAs integrated with up to 16 MS2 binding motifs to enable robust fluorescent signal amplification. These engineered sgRNAs enable multicolor labeling of low-repeat-containing regions using a single sgRNA and of non-repetitive regions with as few as four unique sgRNAs. Then they successfully tracked native chromatin loci throughout the cell cycle and determine differential positioning of transcriptionally active and inactive regions in the nucleus.

Large-scale genetic and epigenetic CRISPR screenings. CRISPR–Cas9-based epigenome editing technologies have enabled precise perturbation of the activity of specific regulatory elements. This technology allows the high-throughput functional annotation of putative regulatory elements in their native chromosomal context.

In 2016, Klann and his colleagues developed a CRISPR–Cas9-based epigenomic regulatory element screening (CERES) for improved high-throughput screening of regulatory element activity in the native genomic context (Klann et al., 2017). Using dCas9KRAB repressor and dCas9p300 activator constructs and lentiviral single guide RNA libraries to target DNase I hypersensitive sites surrounding a gene of interest, we carried out both loss- and gain-of-function screens to identify regulatory elements for the β-globin and HER2 loci in human cells.

References

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709-1712. https://doi.org/10.1126/science.1138140

Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading)*, *151*(Pt 8), 2551-2561. https://doi.org/10.1099/mic.0.28048-0

Cheng, A. W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T. W., Rangarajan, S., Shivalila, C. S., Dadon, D. B., & Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res*, *23*(10), 1163-1171. https://doi.org/10.1038/cr.2013.122

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819-823. https://doi.org/10.1126/science.1231143

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602-607. https://doi.org/10.1038/nature09886

Garneau, J. E., Dupuis, M., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A. H., & Moineau, S. (2010). The CRISPR/Cas bacterial

immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), 67-71. https://doi.org/10.1038/nature09523

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*, *109*(39), E2579-2586. https://doi.org/10.1073/pnas.1208507109

Haft, D. H., Selengut, J., Mongodin, E. F., & Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol*, *1*(6), e60. https://doi.org/10.1371/journal.pcbi.0010060

Holmgaard, A., Alsing, S., Askou, A. L., & Corydon, T. J. (2019). CRISPR Gene Therapy of the Eye: Targeted Knockout of Vegfa in Mouse Retina by Lentiviral Delivery. *Methods Mol Biol*, *1961*, 307-328. https://doi.org/10.1007/978-1-4939-9170-9_19

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., Li, Y., Fine, E. J., Wu, X., Shalem, O., Cradick, T. J., Marraffini, L. A., Bao, G., & Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*, *31*(9), 827-832. https://doi.org/10.1038/nbt.2647

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816-821. https://doi.org/10.1126/science.1225829

Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J. S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res*, *24*(6), 1012-1019. https://doi.org/10.1101/gr.171322.113

Klann, T. S., Black, J. B., Chellappan, M., Safi, A., Song, L., Hilton, I. B., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. (2017). CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat Biotechnol*, *35*(6), 561-568. https://doi.org/10.1038/nbt.3853

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Topkar, V. V., Zheng, Z., & Joung, J. K. (2015). Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol*, *33*(12), 1293-1298. https://doi.org/10.1038/nbt.3404

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T., Zheng, Z., Gonzales, A. P., Li, Z., Peterson, R. T., Yeh, J. R., Aryee, M. J., & Joung, J. K. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, *523*(7561), 481-485. https://doi.org/10.1038/nature14592

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (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. *Biol 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., Yakunin, A. F., van der Oost, J., & Koonin, E.

V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*, *9*(6), 467-477. https://doi.org/10.1038/nrmicro2577

Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., & Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*, *31*(9), 833-838. https://doi.org/10.1038/nbt.2675

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, *322*(5909), 1843-1845. https://doi.org/10.1126/science.1165771

Mojica, F. J., Juez, G., & Rodríguez-Valera, F. (1993). Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. *Mol Microbiol*, *9*(3), 613-621. https://doi.org/10.1111/j.1365-2958.1993.tb01721.x

Mojica, F. J. M., Díez-Villaseñor, C. s., García-Martínez, J., & Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, *60*(2), 174-182. https://doi.org/10.1007/s00239-004-0046-3

Qin, P., Parlak, M., Kuscu, C., Bandaria, J., Mir, M., Szlachta, K., Singh, R., Darzacq, X., Yildiz, A., & Adli, M. (2017). Live cell imaging of low- and non-repetitive chromosome loci using CRISPR-Cas9. *Nat Commun*, *8*, 14725. https://doi.org/10.1038/ncomms14725

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). The Streptococcus thermophilus CRISPR/Cas system provides immunity in

Escherichia coli. *Nucleic Acids Res*, *39*(21), 9275-9282. https://doi.org/10.1093/nar/gkr606

Sorek, R., Kunin, V., & Hugenholtz, P. (2008). CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol*, *6*(3), 181-186. https://doi.org/10.1038/nrmicro1793

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., & Vale, R. D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, *159*(3), 635-646. https://doi.org/10.1016/j.cell.2014.09.039

Thakore, P. I., D'Ippolito, A. M., Song, L., Safi, A., Shivakumar, N. K., Kabadi, A. M., Reddy, T. E., Crawford, G. E., & Gersbach, C. A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*, *12*(12), 1143-1149. https://doi.org/10.1038/nmeth.3630