Origins and Applications of CRISPR-Mediated Genome Editing

John R. Christin, MS,[1,2] and Michael V. Beckert, MS[3]
[1] Ruth L. and David S. Gottesman Institute for Stem Cell and Regenerative Medicine Research, Albert Einstein College of Medicine, Bronx, NY.
[2] Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY.
[3] Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY.

DISCOVERY AND DEVELOPMENT OF CRISPR AND CAS9 (FIGURE 1)

In the late 1980s, a group of researchers interested in the alkaline phosphatase of *Escherichia coli* discovered something odd. In their paper, the authors briefly described a strange genomic topology consisting of a series of 32 nucleotides of unique sequence, flanked by short invariable palindromic repeats on the 3' end of the phosphatase gene that they had been studying (Ishino et al., 1987). The odd genomic architecture they observed is the first known description of a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) array. It would be another 15 years until additional work was done on these novel loci. Further work would reveal that numerous protein-coding genes cluster near CRISPR arrays and that these genes are highly conserved among bacteria and archaea (Jansen et al., 2002). In 2005, a trio of papers began to uncover the function of these pervasive and unusual loci. Bolotin et al., Mojica et al., and Pourcel et al. demonstrated that the unique spacer regions found in CRISPR arrays actually mapped to phage genomes, hinting at CRISPR as a possible adaptive immune response to phage infection though an RNA-guided process (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). The molecular mechanism of this immune response was elucidated in 2012; two papers demonstrated that CRISPR arrays are transcribed into RNA, which is then cleaved and loaded into CRISPR-associated (cas) proteins (cas9, in this case). This RNA:protein complex is sufficient for RNA-guided dsDNA endonuclease activity (Gasiunas et al., 2012; Jinek et al., 2012). Furthermore, Jinek et al. demonstrated that cas9 could be reprogrammed to target novel sequences with an *in vitro* transcribed single guide RNA (sgRNA) (Jinek et al., 2012). This group also demonstrated that two amino acid changes to cas9 could render its nuclease domains non-functional (Jinek et al., 2012), a concept that has been seized upon by other groups to develop novel tools to regulate gene expression.

For many years, researchers had been searching for a tool to induce mutations easily in a targeted fashion. While some headway had been made with Zinc Finger Nucleases, Meganucleases, and Transcription Activator-Like Effector Nucleases (TALENS), all of these techniques had several limitations. Each was either labor intensive, expensive, or both, as the targeting mechanisms were all based on protein-nucleic acid interactions, thereby requiring a custom-designed protein for each gene locus of interest. The promise of RNA-guided nuclease activity afforded by CRISPR-based approaches led numerous groups to recognize immediately this technology's potential to induce targeted, double-stranded breaks (DSB) in eukaryotes, which previously could only be accomplished with much difficulty. DSBs produced by previously available technologies, and now CRISPR-based systems, are repaired by low-fidelity DNA repair pathways, leading to the production of insertion/deletion mutations (indels)—a class of mutations characterized by the random insertion or deletion of nucleotides at the site of the DSB. The introduction of indels into the coding region of a gene can then, either *de novo* or due to a frame shift, introduce a premature stop codon leading to a truncated protein product, or the induction of non-sense mediated decay of the mRNA transcript itself upon expression of the targeted gene. The production of DSBs can also be used to promote the successful knock-in of novel genetic elements by flanking the novel element with homologous sequences derived from the targeted locus, and co-delivering the flanked novel element along with the sgRNA and cas9.

The first demonstration of RNA-guided mutation in eukaryotic cells occurred in 2013 (Cong et al., 2013; Mali et al., 2013). While reprogramming sgRNAs was not a novel discovery at this point—Jennifer Doudna's group had already shown that cas9 could easily be reprogrammed to cleave DNA *in vitro*—these papers were instrumental in providing the scientific community with a well-documented set of tools that could easily be implemented by other labs. Weeks after these papers were published, any lab could obtain CRISPR constructs, purchase a pair of oligonucleotides, perform a simple cloning reaction, and quickly create knock-out or knock-in cell lines (Cong et al., 2013; Mali et al., 2013) or with some additional equipment, animals (Wang et al., 2013). With this single tool, both of these activities have now become technically and financially accessible to a variety of labs, and are no longer confined to the sole domain of industry labs or particularly well-funded academic labs.

However, cutting DNA is by no means the only application for CRISPR. The nuclease-dead cas9 that Doudna's group produced in 2012 was shown to be still capable of binding the targeted locus and disrupting either transcriptional initiation or elongation via steric hindrance, thereby repressing gene expression without inducing DSBs in the genome (Qi et al., 2013). Other groups have further exploited this characteristic by creating cas9 fusion proteins, allowing for fine-tuned adjustment of gene expression (Gilbert et al., 2013), assessing epigenetic state (Hilton et al., 2015; Kearns et al., 2015), and even fluorescent imaging of the genome in live cells (Chen et al., 2013). Furthermore, by taking advantage
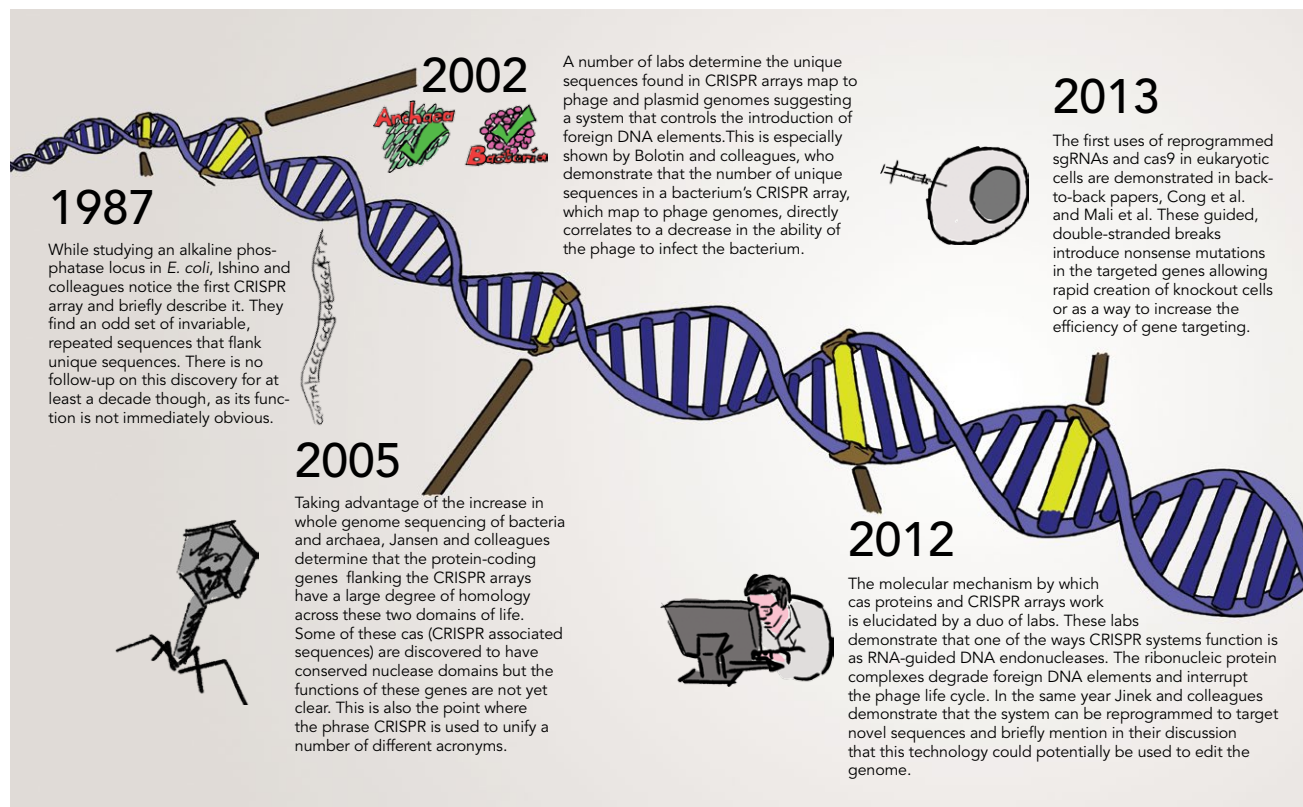
Figure 1 | The Discovery and Development of CRISPR and Cas9

of the modular nature of sgRNAs and the ever-decreasing price of oligonucleotide synthesis, multiple screening libraries (Doench et al., 2016; Joung et al., 2016; Konermann et al., 2015; Sanjana et al., 2014; Shalem et al., 2014; Wang et al., 2014) have been produced to knock out genes by the introduction of indels into the coding sequence, as well as to regulate their expression by targeting the proximal promoters of genes and using some of the fusion proteins described above.

With all these possible uses, CRISPR-based technologies have captured the imagination of biologists, and rightly so. However, with new techniques comes the potential for novel sources of error. Therefore scientists using these systems should consider the following: How specific is the sgRNA in question? Do multiple sgRNAs targeted to the same gene locus recapitulate similar phenotypes? Is cas9 transiently or constitutively expressed? As a control, is cas9 protein used alone or in combination with a non-targeting sgRNA? To paraphrase Voltaire, the perfect experiment is the enemy of the appropriately controlled one. Finally, it is imperative to read and understand the technical details of this powerful technology before implementing it in one's own projects. In order to comprehend fully what other research groups are doing, or when a certain flavor of the technology might be useful for one's own studies, it is essential to develop a familiarity with the capabilities and limitations of CRISPR.

MECHANISMS OF CAS9:SGRNA TARGET BINDING AND DNA REPAIR (FIGURE 2)

In the nucleus, the cas9:sgRNA complex rapidly begins to sample (near the speed of diffusion) the genome for any bases that match its protospacer-adjacent motif (PAM) (Jinek et al., 2014; Knight et al., 2015). Upon recognition of the PAM by cas9 the complex slows down, briefly allowing sgRNA to bind the bases 5' of the PAM. If there is little or no base pairing to the genomic DNA, the complex detaches and samples additional PAMs elsewhere. However, if the sgRNA perfectly matches or nearly perfectly matches the genomic sequence, the sgRNA and its genomic complement enter a central channel of the cas9 protein where the complementary genomic DNA is cleaved by one of the two nuclease domains found in the cas9 protein (Anders et al., 2014; Jinek et al., 2014). Simultaneously, the anti-complementary genomic DNA is fed into a second channel of the of cas9 protein where it is also cleaved (Anders et al., 2014). After cleavage, the cas9:sgRNA complex disassociates from the genomic DNA and continues to sample for additional PAMs (Knight et al., 2015). The presence of a double-stranded break within a cell leads to the activation of DNA damage responses, wherein the cell will either repair the break via template-driven homologous recombination (HR), or by non-homologous end joining (NHEJ), a template-independent repair mechanism. The former method will result in an error-free repair, and if an exogenous template is introduced, this repair method will incorporate novel
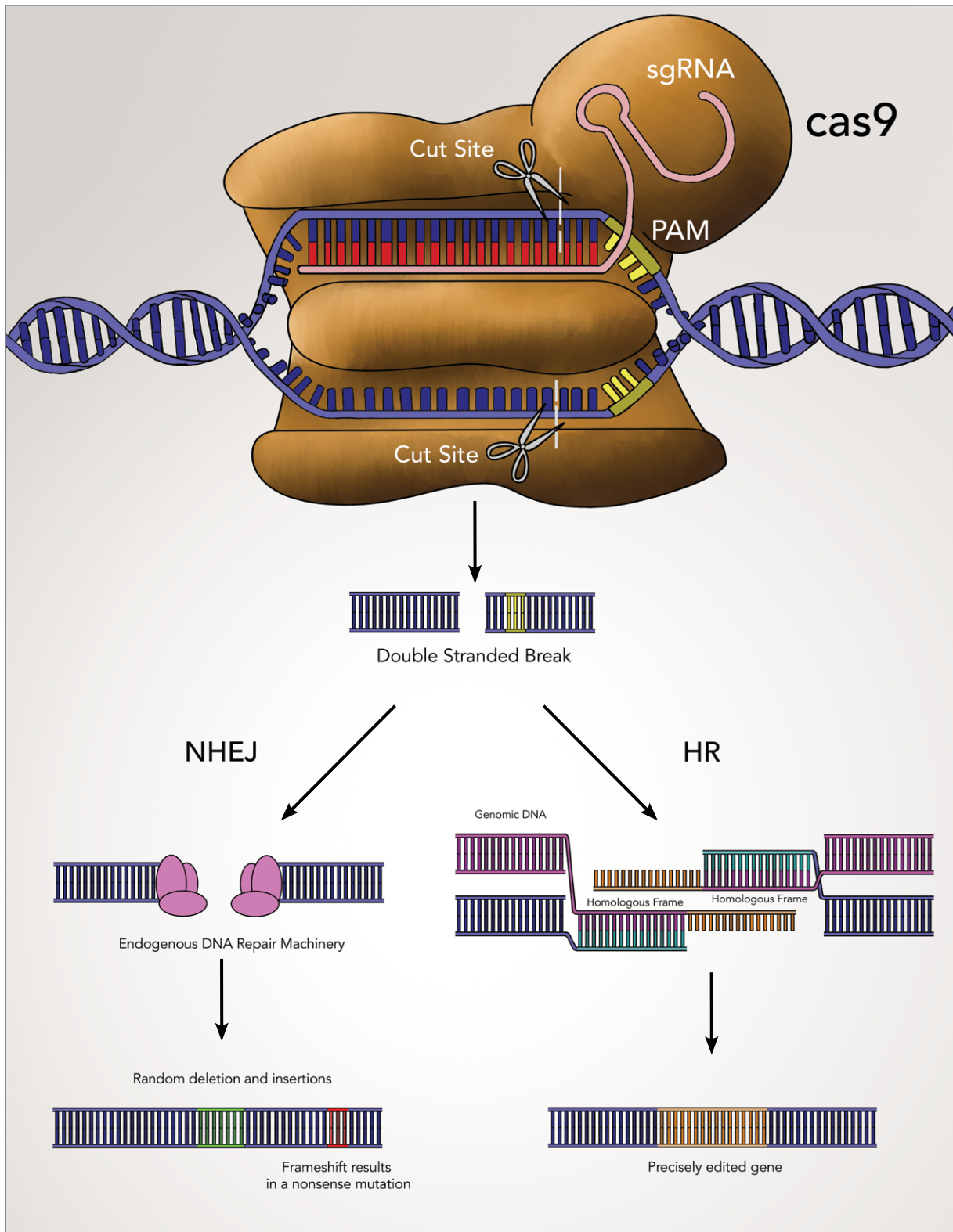
Figure 2 | Mechanisms of cas9:sgRNA Target Binding and DNA Repair

genetic elements into the genome. However, NHEJ will result in the random subtraction or addition of nucleotides at the break site in order to create conditions conducive to sealing the break, but this generally results in a frameshift and a downstream nonsense mutation in the targeted gene.

Corresponding Author: John R. Christin (john.christin@phd.einstein.yu.edu.).

References

Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*, 2551–2561.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., et al. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*, 1479–1491.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*, 819–823.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol*.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, *109*, E2579–E2586.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, *154*, 442–451.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol*, *33*, 510–517.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J. Bacteriol*, *169*, 5429–5433.

Jansen, R., Embden, J.D.A.V., Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol*, *43*, 1565–1575.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*, 816–821.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science, 343*, 1247997.

Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2016). Protocol: Genome-scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening. *bioRxiv*.

Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat Methods, 12*, 401–403.

Knight, S.C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L.B., Bosanac, L., Zhang, E.T., Beheiry, El, M., Masson, J.-B., Dahan, M., et al. (2015). Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*, *350*, 823–826.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature, 517*, 583–588.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science*, *339*, 823–826.

Mojica, F.J.M., Díez-Villaseñor, C.S., 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.

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.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*, 1173–1183.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods*, *11*, 783–784.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, *343*, 84–87.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*,*153*, 910–918.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, *343*, 80–84.