Cell

Leading Edge
Perspective

The Heroes of CRISPR

Eric S. Lander[1,2,3,*]
[1]Broad Institute of MIT and Harvard, 415 Main Street, Cambridge, MA 02142, USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
*Correspondence: lander@broadinstitute.org
http://dx.doi.org/10.1016/j.cell.2015.12.041

Three years ago, scientists reported that CRISPR technology can enable precise and efficient genome editing in living eukaryotic cells. Since then, the method has taken the scientific community by storm, with thousands of labs using it for applications from biomedicine to agriculture. Yet, the preceding 20-year journey—the discovery of a strange microbial repeat sequence; its recognition as an adaptive immune system; its biological characterization; and its repurposing for genome engineering—remains little known. This Perspective aims to fill in this backstory—the history of ideas and the stories of pioneers—and draw lessons about the remarkable ecosystem underlying scientific discovery.

Introduction

It's hard to recall a revolution that has swept biology more swiftly than CRISPR. Just 3 years ago, scientists reported that the CRISPR system—an adaptive immune system used by microbes to defend themselves against invading viruses by recording and targeting their DNA sequences—could be repurposed into a simple and reliable technique for editing, in living cells, the genomes of mammals and other organisms. CRISPR was soon adapted for a vast range of applications—creating complex animal models of human-inherited diseases and cancers; performing genome-wide screens in human cells to pinpoint the genes underlying biological processes; turning specific genes on or off; and genetically modifying plants—and is being used in thousands of labs worldwide. The prospect that CRISPR might be used to modify the human germline has stimulated international debate.

If there are molecular biologists left who have not heard of CRISPR, I have not met them. Yet, if you ask scientists *how* this revolution came to pass, they often have no idea. The immunologist Sir Peter Medawar observed, "The history of science bores most scientists stiff" (Medawar, 1968). Indeed, scientists focus relentlessly on the future. Once a fact is firmly established, the circuitous path that led to its discovery is seen as a distraction.

Yet, the human stories behind scientific advances can teach us a lot about the miraculous ecosystem that drives biomedical progress—about the roles of serendipity and planning, of pure curiosity and practical application, of hypothesis-free and hypothesis-driven science, of individuals and teams, and of fresh perspectives and deep expertise. Such understanding is important for government agencies and foundations that together invest, in the U.S. alone, more than $40 billion in biomedical research. It is also important for a general public who often imagines scientists as lone geniuses cloistered in laboratories. And, for trainees, it is especially valuable to have a realistic picture of scientific careers, as both guide and inspiration.

Over the past several months, I have sought to understand the 20-year backstory behind CRISPR, including the history of ideas and the stories of individuals. This Perspective is based on published papers, personal interviews, and other materials—including rejection letters from journals. At the end, I try to distill some general lessons. (As background, Figure 1 provides a brief overview of a type II CRISPR system, the variety that has been repurposed for genome engineering.)

Most of all, the Perspective describes an inspiring ensemble of a dozen or so scientists who—with their collaborators and other contributors whose stories are not elaborated here—discovered the CRISPR system, unraveled its molecular mechanisms, and repurposed it as a powerful tool for biological research and biomedicine. Together, they are the Heroes of CRISPR.

Discovery of CRISPR

The story starts in the Mediterranean port of Santa Pola on Spain's Costa Blanca, where the beautiful coast and vast salt marshes have for centuries attracted vacationers, flamingoes, and commercial salt producers. (The geography of the story is shown in Figure 2.) Francisco Mojica, who grew up nearby, frequented those beaches, and it was no surprise that, when he began his doctoral studies in 1989 at the University of Alicante, just up the coast, he joined a laboratory working on *Haloferax mediterranei*, an archaeal microbe with extreme salt tolerance that had been isolated from Santa Pola's marshes. His advisor had found that the salt concentration of the growth medium appeared to affect the way in which restriction enzymes cut the microbe's genome, and Mojica set out to characterize the altered fragments. In the first DNA fragment he examined, 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).

The 28-year-old graduate student was captivated and devoted the next decade of his career to unraveling the mystery.

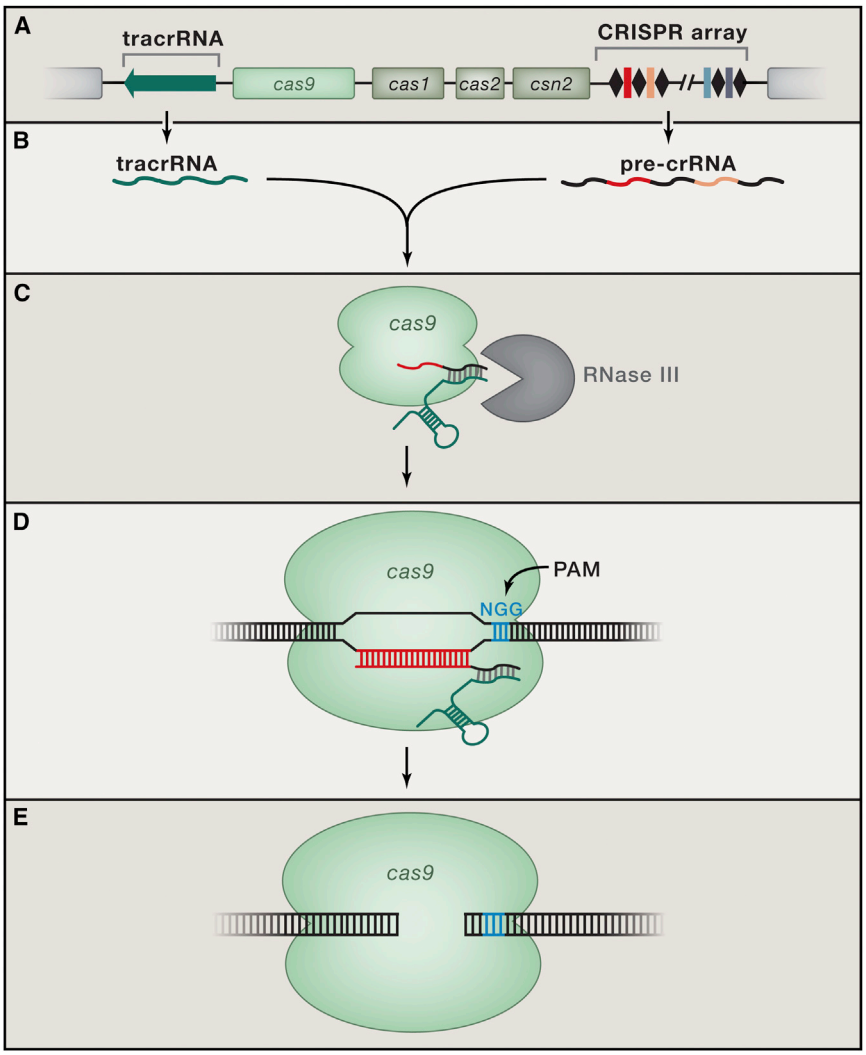

Figure 1. Class 2, Type II CRISPR-Cas9 System from *Streptococcus thermophilus*

Type II systems are the simplest of the three types of CRISPR systems and have been the basis for genome editing technology.

(A) The locus contains a CRISPR array, four protein-coding genes (*cas9*, *cas1*, *cas2*, and *cns2*) and the tracrRNA. The CRISPR array contains repeat regions (black diamonds) separated by spacer regions (colored rectangles) derived from phage and other invading genetic elements. The *cas9* gene encodes a nuclease that confers immunity by cutting invading DNA that matches existing spacers, while the *cas1*, *cas2*, and *cns2* genes encode proteins that function in the acquisition of new spacers from invading DNA.

(B) The CRISPR array and the tracrRNA are transcribed, giving rise to a long pre-crRNA and a tracrRNA.

(C) These two RNAs hybridize via complementary sequences and are processed to shorter forms by Cas9 and RNase III.

(D) The resulting complex (Cas9 + tracrRNA + crRNA) then begins searching for the DNA sequences that match the spacer sequence (shown in red). Binding to the target site also requires the presence of the protospacer adjacent motif (PAM), which functions as a molecular handle for Cas9 to grab on to.

(E) Once Cas9 binds to a target site with a match between the crRNA and the target DNA, it cleaves the DNA three bases upstream of the PAM site. Cas9 contains two endonuclease domains, HNH and RuvC, which cleave, respectively, the complementary and non-complementary strands of the target DNA, creating blunt ends.

He soon discovered similar repeats in the closely related *H. volcanii*, as well as in more distant halophilic archaea. Combing through the scientific literature, he also spotted a connection with eubacteria: a paper by a Japanese group (Ishino et al., 1987) had mentioned a repeat sequence in *Escherichia coli* that had a similar structure, although no sequence similarity, to the *Haloferax* repeats. These authors had made little of the observation, but Mojica realized that the presence of such similar structures in such distant microbes must signal an important function in prokaryotes. He wrote up a paper reporting this new class of repeats (Mojica et al., 1995) before heading off for a short post-doctoral stint at Oxford.

Mojica returned home to take up a faculty position at the University of Alicante. Because the school had hardly any start-up funds or lab space, he turned to bioinformatics to investigate the strange repeats, which he dubbed short regularly spaced repeats (SRSRs); the name would later be changed, at his suggestion, to clustered regularly interspaced palindromic repeats (CRISPR) (Jansen et al., 2002; Mojica and Garrett, 2012).

By 2000, Mojica had found CRISPR loci in 20 different microbes—including *Mycobacterium tuberculosis*, *Clostridium difficile*, and the plague bacteria *Yersinia pestis* (Mojica et al., 2000). Within 2 years, researchers had doubled the census and cataloged key features of loci—including the presence of specific CRISPR-associated (cas) genes in the immediate vicinity, which were presumably related to their function (Jansen et al., 2002). (Table 1 summarizes the modern classification of CRISPR systems.)

But what was the function of the CRISPR system? Hypotheses abounded: it was variously proposed to be involved in gene regulation, replicon partitioning, DNA repair, and other roles (Mojica and Garrett, 2012). But most of these guesses rested on little or no evidence, and one by one they proved to be wrong. As with the discovery of CRISPR, the critical insight came from bioinformatics.

CRISPR Is an Adaptive Immune System

During the August holiday in 2003, Mojica escaped the scorching heat of Santa Pola's beaches and took refuge in his air-conditioned office in Alicante. By now the clear leader in the nascent CRISPR field, he had turned his focus from the repeats themselves to the spacers that separated them. Using his word processor, Mojica painstakingly extracted each spacer and inserted it into the BLAST program to search for similarity with any other

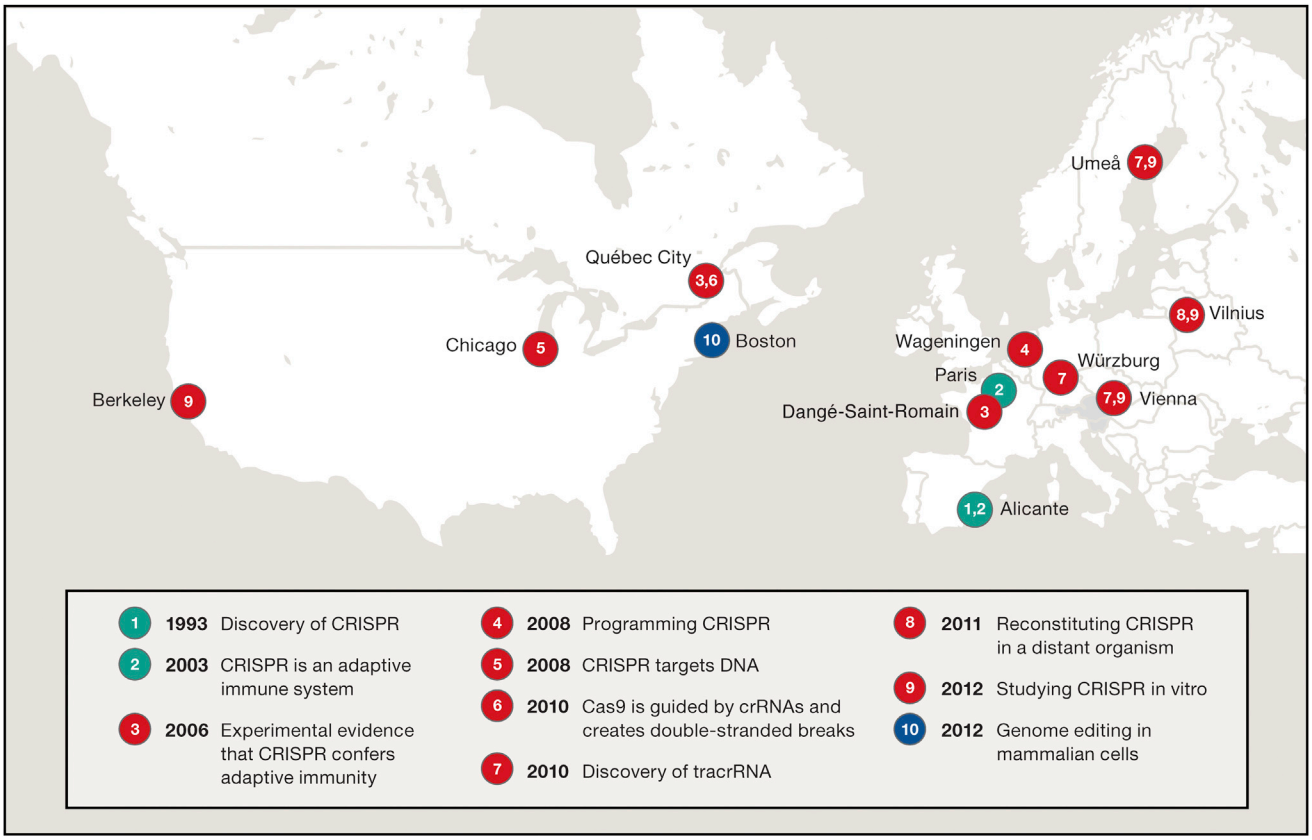

Figure 2. The Twenty-Year Story of CRISPR Unfolded across Twelve Cities in Nine Countries
For each "chapter" in the CRISPR "story," the map shows the sites where the primary work occurred and the first submission dates of the papers. Green circles refer to the early discovery of the CRISPR system and its function; red to the genetic, molecular biological, and biochemical characterization; and blue to the final step of biological engineering to enable genome editing.

known DNA sequence. He had tried this exercise before without success, but the DNA sequence databases were continually expanding and this time he struck gold. In a CRISPR locus that he had recently sequenced from an *E. coli* strain, one of the spacers matched the sequence of a P1 phage that infected many *E. coli* strains. However, the particular strain carrying the spacer was known to be resistant to P1 infection. By the end of the week, he had slogged through 4,500 spacers. Of 88 spacers with similarity to known sequences, two-thirds matched viruses or conjugative plasmids related to the microbe carrying the spacer. Mojica realized that CRISPR loci must encode the instructions for an adaptive immune system that protected microbes against specific infections.

Mojica went out to celebrate with colleagues over cognac and returned the next morning to draft a paper. So began an 18-month odyssey of frustration. Recognizing the importance of the discovery, Mojica sent the paper to *Nature*. In November 2003, the journal rejected the paper without seeking external review; inexplicably, the editor claimed the key idea was already known. In January 2004, the *Proceedings of the National Academy of Sciences* decided that the paper lacked sufficient "novelty and importance" to justify sending it out to review. *Molecular Microbiology* and *Nucleic Acid Research* rejected the paper in turn. By now desperate and afraid of being scooped, Mojica sent the paper to *Journal of Molecular Evolution*. After 12 more months of review and revision, the paper reporting CRISPR's likely function finally appeared on February 1, 2005 (Mojica et al., 2005).

At about the same time, CRISPR was the focus of attention in another, rather unlikely, venue: a unit of the French Ministry of Defense, some 30 miles south of Paris. Gilles Vergnaud, a human geneticist trained at the Institut Pasteur, had received doctoral and post-doctoral support from the Direction Générale de l'Armement. When he completed his studies in 1987, he joined the government agency to set up its first molecular biology lab. For the next 10 years, Vergnaud continued his work on human genetics. But when intelligence reports in the late 1990s raised concerns that Saddam Hussein's regime in Iraq was developing biological weapons, the Ministry of Defense asked Vergnaud in 1997 to shift his group's focus to forensic microbiology—developing methods to trace the source of pathogens based on subtle genetic differences among strains. Establishing a joint lab with the nearby Institute of Genetics and Microbiology at Université Paris-Sud, he set out to use tandem-repeat polymorphisms—which were the workhorse of forensic DNA fingerprinting in humans—to characterize strains of the bacteria responsible for anthrax and plague.

Table 1. Classification and Examples of CRISPR Systems

Class	Type	Subtype	Hallmarks	Example effector	Example organism	Studies Cited
Class 1	Type I		multisubunit effector complex; Cas3	Cascade	*E. coli*	Brouns et al., 2008
	Type III	III-A	multisubunit effector complex; Csm effector module; DNA targeting	Cas10-Csm	*S. epidermidis*	Marraffini and Sontheimer, 2008
		III-B	multisubunit effector complex; Cmr effector module; RNA targeting	Cmr	*P. furiosus*	Hale et al., 2009
Class 2	Type II		single protein effector; tracrRNA	Cas9	*S. thermophilus*	Bolotin et al., 2005; Barrangou et al., 2007; Sapranauskas et al., 2011; Gasiunas et al., 2012
					S. pyogenes	Deltcheva et al., 2011; Jinek et al., 2012; Cong et al., 2013; Mali et al., 2013
	Type V		single protein effector; single-RNA guided	Cpf1	*F. novicida*	Zetsche et al., 2015

CRISPR systems are currently organized into two overarching classes: Class 1, which contain multi-subunit effectors, and Class 2, which contain single protein effectors. These classes are subdivided into five types (Makarova et al., 2015), with type IV remaining a putative type within Class 1. Although only Class 2 systems have been adapted for genome engineering, the results described in this review emerged from studying a diversity of CRISPR-Cas systems. (Type III-B systems are not discussed but represent an unusual system that targets RNA rather than DNA [Hale et al., 2009].)

The French Defense Ministry had access to a unique trove of 61 *Y. pestis* samples from a plague outbreak in Vietnam in 1964–1966. Vergnaud found that these closely related isolates were identical at their tandem-repeat loci—with a sole exception of a site that his colleague Christine Pourcel discovered was the CRISPR locus. The strains occasionally differed by the presence of new spacers, which were invariably acquired in a polarized fashion at the "front" end of the CRISPR locus (Pourcel et al., 2005). Strikingly, many of the new spacers corresponded to a prophage present in the *Y. pestis* genome. The authors proposed that the CRISPR locus serves in a defense mechanism—as they put it, poetically, "CRISPRs may represent a memory of 'past genetic aggressions.'" Vergnaud's efforts to publish their findings met the same resistance as Mojica's. The paper was rejected from the *Proceedings of the National Academy of Sciences*, *Journal of Bacteriology, Nucleic Acids Research*, and *Genome Research*, before being published in *Microbiology* on March 1, 2005.

Finally, a third researcher—Alexander Bolotin, a Russian émigré who was a microbiologist at the French National Institute for Agricultural Research—also published a paper describing the extrachromosomal origin of CRISPR, in *Microbiology* in September 2005 (Bolotin et al., 2005). His report was actually submitted a month after Mojica's February 2005 paper had already appeared—because his submission to another journal had been rejected. Notably, Bolotin was the first to speculate how CRISPR conferred immunity—proposing that transcripts from the CRISPR locus worked by anti-sense RNA inhibition of phage gene expression. Although reasonable, the guess would prove to be wrong.

Experimental Evidence that CRISPR Confers Adaptive Immunity and Employs a Nuclease

Like Mojica, Philippe Horvath could hardly have chosen a thesis topic that was more local or less sexy. As a Ph.D. student at the University of Strasbourg, he concentrated on the genetics of a lactic-acid bacteria used in the production of sauerkraut—the central ingredient in the Alsatian specialty *choucroute garnie*. Given his interest in food science, Horvath skipped doing postdoctoral research and in late 2000 joined Rhodia Food, a maker of bacterial starter cultures located in Dangé-Saint-Romain in western France, to set up its first molecular biology lab. The company was later acquired by the Danish firm Danisco, which was itself acquired by DuPont in 2011.

Rhodia Food was interested in Horvath's microbiological skills because other lactic-acid bacteria, such as *Streptococcus thermophilus*, are used to make dairy products, such as yogurt and cheese. Horvath's mission included developing DNA-based methods for precise identification of bacterial strains and overcoming the frequent phage infections that plagued industrial cultures used in dairy fermentation. Understanding how certain *S. thermophilus* strains protect themselves from phage attack was thus of both scientific interest and economic importance.

After learning about CRISPR at a Dutch conference on lactic-acid bacteria in late 2002, Horvath began using it to genotype his strains. By late 2004, he noticed a clear correlation between spacers and phage resistance—as would be reported just a few months later by Mojica and Vergnaud. In 2005, Horvath and colleagues—including Rodolphe Barrangou, a newly minted Ph.D. at Danisco USA, and Sylvain Moineau, a distinguished phage biologist at Université Laval in Québec City—set out to directly test the hypothesis that CRISPR was an adaptive immune system. Notably, Moineau had also been an industrial scientist. He had earned his Ph.D in Food Sciences at Laval, also studying lactic-acid bacteria, and had worked at Unilever Corporation before returning to academia at Laval; he had been collaborating with Rhodia Food since 2000.

Using a well-characterized phage-sensitive *S. thermophilus* strain and two bacteriophages, these investigators performed genetic selections to isolate phage-resistant bacteria. Rather than harboring classical resistance mutations (such as in a cell-surface receptor required for phage entry), the resistant

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.

They also studied the role of two of the *cas* genes: *cas7* and *cas9*. Bacteria required *cas7* in order to gain resistance, but those carrying a phage-derived spacer did not need the gene to remain resistant—suggesting that Cas7 was involved in generating new spacers and repeats, but not in immunity itself. In contrast, *cas9*—whose sequence contained two types of nuclease motifs (HNH and RuvC) and whose product thus presumably cut nucleic acids (Bolotin et al., 2005; Makarova et al., 2006)—was necessary for phage resistance; the Cas9 protein was an active component of the bacterial immune system. (Warning: In the early CRISPR literature, the now-famous *cas9* gene was called *cas5* or *csn1*.)

Finally, they found that rare phage isolates that overcame CRISPR-based immunity carried single-base changes in their genomes that altered the sequence corresponding to the spacers. Immunity thus depended on a precise DNA sequence match between spacer and target.

Programming CRISPR

John van der Oost, who got his Ph.D. from the Free University of Amsterdam in 1989, originally set out to solve the world's clean-energy needs by using cyanobacteria to produce biofuels. He studied metabolic pathways in bacteria, working in Helsinki and Heidelberg before returning to Amsterdam. In 1995, Wageningen University offered him a permanent position—but with a catch: they wanted him to expand a group working on extremophile microbes. van der Oost, who had once heard a talk while in Germany about *Sulfolobus solfataricus*, which thrives in the hot springs of Yellowstone National Park, was game to investigate the evolutionary differences in the metabolic pathways of these strange microbes. He began to collaborate with Eugene Koonin—an expert in microbial evolution and computational biology at the National Center for Biotechnology Information (NCBI) at the National Institutes of Health. Koonin had begun working on classifying and analyzing CRISPR systems, and on a visit in 2005, he introduced van der Oost to the then-obscure field of CRISPR (Makarova et al., 2006).

van der Oost had just received a major grant from the Dutch National Science Foundation. In addition to working on the problem described in his proposal, he decided to use some of the funding to study CRISPR. (In his report to the agency 5 years later, he underscored the value of the agency's policy of allowing researchers the freedom to shift their scientific plans.)

He and his colleagues inserted an *E. coli* CRISPR system into another *E. coli* strain that lacked its own endogenous system. This allowed them to biochemically characterize a complex of five Cas proteins, termed Cascade (Brouns et al., 2008). (*E. coli* has the more complex Class 1, type I CRISPR system, in which the functions of Cas9 are instead performed by the Cascade complex, together with the nuclease Cas3. See Table 1.)

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 that co-purified with 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. This finding supported 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 create the first artificial CRISPR arrays—programming CRISPR to target four essential genes in lambda (λ) phage. As they predicted, the strains carrying the new CRISPR sequence showed resistance to phage λ. It was the first case of directly programming CRISPR-based immunity—a flu shot for bacteria.

The results also hinted that the target of CRISPR was not RNA (as Bolotin had proposed) but, rather, DNA. The authors had designed two versions of the CRISPR array—one in the anti-sense direction (complementary to both the mRNA and coding strand of the DNA locus) and one in the sense direction (complementary only to the other DNA strand). Although the spacers varied in their efficacy, the fact that the sense version worked strongly suggested that the target was not mRNA. Still, the evidence was indirect. With the journal editors urging caution about drawing a firm conclusion, van der Oost's paper in *Science* offered the notion that CRISPR targets DNA as a "hypothesis."

CRISPR Targets DNA

Luciano Marraffini was finishing his Ph.D., working on *Staphylococcus*, at the University of Chicago, when he learned about CRISPR from Malcolm Casadaban, a faculty member in the department who was a world authority on phage genetics. Casadaban had immediately seen the importance of discovery in 2005 that CRISPR was likely to be an adaptive immune system and talked about CRISPR to everyone who would listen. Like many in the phage community, Marraffini was convinced that CRISPR could not work by RNA interference because this mechanism would be too inefficient to overcome the explosive growth that occurs upon phage infection. Instead, he reasoned, CRISPR must cut DNA—functioning, in effect, like a restriction enzyme.

Marraffini was eager to pursue post-doctoral work in one of the handful of groups in the world studying CRISPR, but his wife had a good job as a translator in the Cook County, Illinois criminal courts and he felt he should stay in Chicago. He persuaded Erik Sontheimer, a biochemist at Northwestern University who had been working on RNA splicing and RNA interference, to let him join his lab to work on CRISPR.

Even before moving to Northwestern, Marraffini started working on CRISPR even as he completed his graduate work—exploring whether the *Staphylococcus* CRISPR system could block plasmid conjugation. He noticed that a strain of *Staphylococcus epidermidis* had a spacer that matches a region of the *nickase* (*nes*) gene encoded on plasmids from antibiotic-resistant *Staphylococcus aureus*. He showed that these plasmids cannot be transferred to *S. epidermidis* but that disrupting either the *nes* sequence in the plasmid or the matching spacer sequences in the CRISPR locus in the genome abolishes interference (Marraffini and Sontheimer, 2008). Clearly, CRISPR blocked the plasmids, just as it blocked viruses.

Marraffini and Sontheimer thought briefly about trying to reconstitute the CRISPR system in vitro to demonstrate that it cuts DNA. But the *S. epidermidis* system was too complicated—it had nine *cas* genes—and it was still too poorly characterized. Instead, they turned to molecular biology. 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 and Sontheimer recognized that CRISPR was essentially a programmable restriction enzyme. Their paper was the first to explicitly predict that CRISPR might be repurposed for genome editing in heterologous systems. "From a practical standpoint," they declared, "the ability to direct the specific addressable destruction of DNA that contains any given 24- to 48-nucleotide target sequence could have considerable functional utility, especially if the system can function outside of its native bacterial or archaeal context." They even filed a patent application including the use of CRISPR to cut or correct genomic loci in eukaryotic cells, but it lacked sufficient experimental demonstration and they eventually abandoned it (Sontheimer and Marraffini, 2008).

Cas9 Is Guided by crRNAs and Creates Double-Stranded Breaks in DNA

Following the seminal study in 2007 confirming that CRISPR is an adaptive immune system (Barrangou et al., 2007), Sylvain Moineau continued to collaborate with Danisco to understand the mechanism by which CRISPR cleaves DNA.

The problem was that CRISPR was normally so efficient that Moineau and his colleagues could not readily observe how invading DNA was destroyed. However, they caught a lucky break while studying plasmid interference in *S. thermophilus*. The investigators found a handful of bacterial strains in which 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. Consistent with their earlier results (Barrangou et al., 2007), they showed that 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, a key sequence feature whose function they had characterized in earlier papers (Deveau et al., 2008; Horvath et al., 2008). 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.

Discovery of tracrRNA

Despite intense study of the CRISPR-Cas9 system, one additional piece of the puzzle was missing—a small RNA that would come to be called *trans*-activating CRISPR RNA (tracrRNA). In fact, the discoverers, Emmanuelle Charpentier and Jörg Vogel, were not specifically looking to study the CRISPR system; they were simply trying to identify microbial RNAs.

Charpentier had earned her Ph.D. in microbiology from Pasteur Institute in 1995 and did post-doctoral work in New York for 6 years before starting her own lab at the University of Vienna in 2002 and Umeå, Sweden in 2008. After discovering an unusual RNA that controls virulence in *Streptococcus pyogenes* (Mangold et al., 2004), she became interested in identifying additional regulatory RNAs in microbes. She used bioinformatics programs to scan intergenic regions in *S. pyogenes* for structures, suggesting that they might encode non-coding RNAs. She had found several candidate regions—including one near the CRISPR locus—but they were hard to follow up without direct information about the RNAs themselves.

The solution appeared when Charpentier met Vogel at the 2007 meeting of RNA Society in Madison, Wisconsin. Trained as a microbiologist in Germany, Vogel had begun focusing on finding RNAs in pathogens during his postdoctoral work in Uppsala and Jerusalem and had continued this work when he started his own group in 2004 at the Max Planck Institute for Infection Biology in Berlin. (Five years later, he would move to Würzburg to lead a research center on infectious disease.) With the recent advent of "next-generation sequencing" technology, Vogel realized that massively parallel sequencing would make it possible to produce comprehensive catalogs of any microbial transcriptome. He had just applied the approach to *Helicobacter pylori*, the bacterium responsible for stomach ulcers (Sharma et al., 2010), and was working on various other bugs. Charpentier and Vogel decided to turn the shotgun on *S. pyogenes* as well.

The approach yielded a striking result: the third-most abundant class of transcript—after only ribosomal RNA and transfer RNA—was a 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).

Later studies would reveal that tracrRNA also has another key role. 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 (Jinek et al., 2012; Siksnys et al., 2012).

Reconstituting CRISPR in a Distant Organism

Virginijus Siksnys grew up in Soviet-era Lithuania and graduated from Vilnius University before leaving home in the early 1980s to get a Ph.D. at Moscow State University, where he studied enzyme kinetics. When he returned home to Vilnius, he joined the Institute of Applied Enzymology to study the then-hot field

of restriction enzymes. After two decades, though, he was bored with characterizing restriction enzymes. Horvath, Barrangou, and Moineau's 2007 paper re-ignited his fascination with bacterial barriers to foreign DNA. As a chemist, he felt that he would only understand CRISPR if he could reconstitute it in vitro.

His first step was to test whether he had all of the necessary components. He and his collaborators set out to see whether the CRISPR system from *S. thermophilus* could be reconstituted in fully functional form in a very distant microbe, *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.

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 known. The system had been completely dissected 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.

Studying CRISPR In Vitro

Using their heterologous expression system in *E. coli*, Siksnys and his colleagues purified the *S. thermophilus* Cas9-crRNA complex by using a streptavidin tag on Cas9 and studied its activity 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. (They would ultimately drop the second reconstitution from their revised paper, but they reported all of the work in their published U.S. patent application filed in March 2012 [Siksnys et al., 2012]).

Around the same time, Charpentier had begun biochemical characterization of CRISPR with a colleague in Vienna. When she lectured about tracrRNA at an American Society for Microbiology meeting in Puerto Rico in March 2011, she met Jennifer Doudna—a world-renowned structural biologist and RNA expert at the University of California, Berkeley. After growing up in Hawaii, Doudna had received her Ph.D. at Harvard, working with Jack Szostak to re-engineer an RNA self-splicing intron into a ribozyme capable of copying an RNA template, and had then done postdoctoral work with Tom Cech at the University of Colorado, where she had solved crystal structures of ribozymes. In her own lab (first at Yale in 1994 and then at Berkeley starting in 2002), she characterized RNA-protein complexes underlying diverse phenomena, such as internal ribosome entry sites and processing of microRNAs. She had been using crystallography and cryo-electron microscopy to solve structures of components of the Cascade complex of type I CRISPR systems, the more complex systems used in microbes such as *E. coli*.

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.

Siksnys submitted his paper to *Cell* on April 6, 2012. Six days later, the journal rejected the paper without external review. (In hindsight, *Cell*'s editor agrees the paper turned out to be very important.) Siksnys condensed the manuscript and sent it on May 21 to the *Proceedings of the National Academy of Sciences*, which published it online on September 4. Charpentier and Doudna's paper fared better. Submitted *to Science* 2 months after Siksnys's on June 8, it sailed through review and appeared online on June 28.

Both groups clearly recognized the potential for biotechnology, with Siksnys declaring that "these findings pave the way for engineering of universal programmable RNA-guided DNA endonucleases," and Charpentier and Doudna noting "the potential to exploit the system for RNA-programmable genome editing." (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.)

Genome Editing in Mammalian Cells

When scientists in the late 1980s devised a way to alter mammalian genomes in living cells, it transformed biomedical research—including making it possible to insert DNA at a specific location in mouse embryonic stem cells and then produce mice carrying the genetic modification (reviewed in Capecchi, 2005). While revolutionary, the process was inefficient—requiring selection and screening to identify the one-in-a-million cells in which homologous recombination had swapped a gene with a modified version supplied by the experimenter. In the mid-1990s, mammalian biologists—building on observations by yeast geneticists—found that introducing a double-stranded break at a genomic locus (using a "meganuclease," an endonuclease with an extremely rare recognition site) dramatically increased the frequency of homologous recombination, as well as small deletions caused by non-homologous end joining (reviewed in Haber, 2000 and Jasin and Rothstein, 2013). The secret to efficient genome editing, they realized, was to find a reliable method to produce a double-stranded break at any desired location. The first general strategy was to use zinc-finger nucleases (ZFNs)—fusion proteins composed of a zinc-finger DNA-binding

domain and a DNA-cleavage domain, taken from a restriction enzyme, that bind and cut a genomic locus (Bibikova et al., 2001). Scientists soon demonstrated the use of ZFNs for site-specific gene editing by homologous recombination in the fruit fly and mouse (Bibikova et al., 2003; Porteus and Baltimore, 2003). By 2005, a group at Sangamo Biosciences reported successful correction of a mutation in the gene for severe combined immune syndrome in a human cell line (Urnov et al., 2005). However, fashioning ZFNs that reliably recognize specific sites proved to be slow and finicky. A better solution emerged after two groups described in late 2009 a remarkable class of transcription-activating proteins called TALEs, from the plant pathogen *Xanthomonas* (Boch et al., 2009; Moscou and Bogdanove, 2009), which use a precise code of modular domains to target specific DNA sequences. Still, the approach entailed considerable work, requiring a new protein for each target.

Since Marraffini and Sontheimer's 2008 paper showing that CRISPR was a programmable restriction enzyme, researchers had grasped that CRISPR might provide a powerful tool for cutting, and thereby editing, specific genomic loci—*if* it could be made to work in mammalian cells. But this was a big "if." In contrast to microbes, mammalian cells have very different internal environments and their genomes are 1,000-fold larger, reside in nuclei, and are embedded in an elaborate chromatin structure. Attempts to transfer other simple microbial systems, such as self-splicing group II introns, had failed, and efforts to use nucleic acids to target genomic loci had been problematic. Could CRISPR be re-engineered to become a robust system for editing the human genome? As late as September 2012, experts were skeptical (Barrangou 2012; Carroll, 2012).

Feng Zhang moved at age 11 from Shijiazhuang, China to Des Moines, Iowa. He got hooked on molecular biology at a Saturday enrichment course and, by age 16, was working 20 hours a week in a local gene-therapy lab. As a Harvard undergraduate, he became interested in the brain when a classmate was stricken by severe depression, and he later pursued a Ph.D. in chemistry at Stanford with neurobiologist and psychiatrist Karl Deisseroth, where they (together with Edward Boyden) developed optogenetics—a revolutionary technique whereby neurons carrying a microbial light-dependent channel protein can be triggered to fire by light pulses. As an independent investigator in Boston (first as a Junior Fellow at Harvard and then at MIT's Department of Brain and Cognitive Sciences and the Broad Institute), Zhang aimed to further expand the molecular toolbox for studying neurobiology. After developing a way to use light to activate gene expression (by coupling a DNA-binding domain and a transcription-activation domain to two plant proteins that bind each other in the presence of light), he began searching for a general way to program transcription factors. When TALEs were deciphered, Zhang, with his collaborators Paola Arlotta and George Church (and, independently, a group from Sangamo BioSciences), successfully repurposed them for mammals—making it possible to activate, repress, or edit genes with precision (Zhang et al., 2011; Miller et al., 2011). Still, he remained on the lookout for a better approach.

In February 2011, Zhang heard a talk about CRISPR from Michael Gilmore, a Harvard microbiologist, and was instantly captivated. He flew the next day to a scientific meeting in Miami but remained holed up in his hotel room digesting the entire CRISPR literature. When he returned, he set out to create a version of *S. thermophilus* Cas9 for use in human cells (with optimized codons and a nuclear-localization signal). By April 2011, he had found that, by expressing Cas9 and an engineered CRISPR RNA targeting a plasmid carrying a luciferase gene, he could decrease luminescence levels in human embryonic kidney cells. Still, the effect was modest.

Zhang spent the next year optimizing the system. He explored ways to increase the proportion of Cas9 that went to the nucleus. When he discovered that *S. thermophilus* Cas9 was unevenly distributed within the nucleus (it clumped in the nucleolus), he tested alternatives and found that *S. pyogenes* Cas9 was much better distributed. He found that mammalian cells, though lacking microbial RNaseIII, could still process crRNA, albeit differently than in bacteria. He tested various isoforms of tracrRNA to identify one that was stable in human 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. Moreover, multiple genes could be edited simultaneously by programming the CRISPR arrays with spacers matching each. When Charpentier and Doudna's paper appeared in early summer, 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 3′ hairpin solved the problem (Cong et al., 2013; Zhang, 2012). (Zhang would soon go on to show that CRISPR was even more versatile: it could be used to create complex mouse models of inherited diseases and somatic cancer in weeks and to perform genome-wide screening to find the essential genes in a biological process—and it could be made more accurate by decreasing "off-target" cutting. He and Koonin, the computational biologist who had worked with van der Oost, would also find new Class 2 CRISPR systems, including one with a nuclease that cuts differently than Cas9 and requires only crRNA without tracrRNA [Zetsche et al., 2015]). Zhang submitted a paper reporting mammalian genome editing on October 5, 2012, which appeared in *Science* on January 3, 2013 (Cong et al., 2013); it would become the most cited paper in the field, with his reagents being distributed by the non-profit organization Addgene in response to more than 25,000 requests over the next 3 years.

About a month later, on October 26, George Church, a brilliant—and colorful—senior Harvard professor with deep expertise in genomics and synthetic biology who had collaborated with Zhang, submitted a paper on genome editing in human cells. Since his time as a graduate student with DNA-sequencing pioneer Walter Gilbert at Harvard in the late 1970s, Church had focused on developing powerful technologies for "reading" and "writing" genomes at large scale—as well as stirring societal debate with provocative proposals, such as to use synthetic biology to revive wooly mammoths and Neanderthals. Aware

of Zhang's efforts and stimulated by Charpentier and Doudna's paper, Church set out to test crRNA-tracrRNA fusions in mammalian cells. Like Zhang, he found that short fusions were inefficient in vivo but that full-length fusions worked well. He targeted seven sites and demonstrated both non-homologous end-joining and homologous recombination. His paper appeared back-to-back with Zhang's (Mali et al., 2013). (Church and others would soon use CRISPR to create improved "gene drives"—synthetic genes able to spread rapidly through natural populations—raising excitement about applications like fighting malaria-carrying mosquitos and worries about disrupting ecosystems. He would also seek to facilitate pig-to-human transplants by using CRISPR to inactivate retroviruses in the porcine genome.)

By late summer 2012—with the in vitro studies gaining attention and news of successful in vivo genome editing spreading before publication—several other groups were racing to perform proof of principle experiments demonstrating genome cleavage, albeit not editing. Doudna, with assistance from Church, submitted a paper demonstrating low-level cutting at one genomic site (Pandika 2014; Jinek et al., 2013). Jin-Soo Kim, a professor at Korea's Seoul National University who had worked on genome editing with ZFNs and TALEs, showed cutting at two sites (Cho et al., 2013). In both cases, the cleavage was inefficient because the sgRNAs lacked the critical 3′ hairpin of tracrRNA. Keith Joung, a Harvard professor who had also been a leader in using ZFNs and TALEs for genome editing, went further. Using the full-length sgRNA structure provided by his collaborator Church, Joung established through experiments in zebrafish that CRISPR could be used to efficiently produce deletions in the germline (Hwang et al., 2013). These short papers, submitted in late 2012 and accepted soon after Zhang's and Church's papers were published in early January 2013, appeared online in late January.

CRISPR Goes Viral

In early 2013, Google searches for "CRISPR" began to skyrocket—a trend that has continued unabated. Within a year, investigators had reported the use of CRISPR-based genome editing in many organisms—including yeast, nematode, fruit fly, zebrafish, mouse, and monkey. Scientific and commercial interest in potential applications in human therapeutics and commercial agriculture began to heat up—as did social concerns about the prospect that the technology could be used to produce designer babies.

The early pioneers of CRISPR continued to push the frontiers, but they were no longer alone. Scientists around the world poured in—a new cadre of heroes who further elucidated the biology of CRISPR, improved and extended the technology for genome editing, and applied it to a vast range of biological problems. It is impossible within the bounds of this Perspective to do justice to these contributions; the reader is referred to recent reviews (Barrangou and Marraffini, 2014; Hsu et al., 2014; van der Oost et al., 2014; Sander and Joung, 2014; Jiang and Marraffini, 2015; Sternberg and Doudna, 2015; Wright et al., 2016).

The once-obscure microbial system—discovered 20 years earlier in a Spanish salt marsh—was now the focus of special issues of scientific journals, headlines in the *New York Times*, biotech start-ups, and international ethics summits (Travis, 2015). CRISPR had arrived.

The Lessons of CRISPR

The story of CRISPR is rich with lessons about the human ecosystem that produces scientific advances, with relevance to funding agencies, the general public, and aspiring researchers.

The most important is that medical breakthroughs often emerge from completely unpredictable origins. The early heroes of CRISPR were not on a quest to edit the human genome—or even to study human disease. Their motivations were a mix of personal curiosity (to understand bizarre repeat sequences in salt-tolerant microbes), military exigency (to defend against biological warfare), and industrial application (to improve yogurt production).

The history also illustrates the growing role in biology of "hypothesis-free" discovery based on big data. The discovery of the CRISPR loci, their biological function, and the tracrRNA all emerged not from wet-bench experiments but from open-ended bioinformatic exploration of large-scale, often public, genomic datasets. "Hypothesis-driven" science of course remains essential, but the 21st century will see an increasing partnership between these two approaches.

It is instructive that so many of the Heroes of CRISPR did their seminal work near the very start of their scientific careers (including Mojica, Horvath, Marraffini, Charpentier, Vogel, and Zhang)—in several cases, before the age of 30. With youth often comes a willingness to take risks—on uncharted directions and seemingly obscure questions—and a drive to succeed. It's an important reminder at a time that the median age for first grants from the NIH has crept up to 42.

Notably, too, many did their landmark work in places that some might regard as off the beaten path of science (Alicante, Spain; France's Ministry of Defense; Danisco's corporate labs; and Vilnius, Lithuania). And, their seminal papers were often rejected by leading journals—appearing only after considerable delay and in less prominent venues. These observations may not be a coincidence: the settings may have afforded greater freedom to pursue less trendy topics but less support about how to overcome skepticism by journals and reviewers.

Finally, the narrative underscores that scientific breakthroughs are rarely eureka moments. They are typically ensemble acts, played out over a decade or more, in which the cast becomes part of something greater than what any one of them could do alone. It's a wonderful lesson for the general public, as well as for a young person contemplating a life in science.

REFERENCES

Barrangou, R. (2012). RNA-mediated programmable DNA cleavage. Nat. Biotechnol. *30*, 836–838.

Barrangou, R., and Marraffini, L.A. (2014). CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. Mol. Cell *54*, 234–244.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.G., and Chandrasegaran, S. (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol. Cell. Biol. *21*, 289–297.

Bibikova, M., Beumer, K., Trautman, J.K., and Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. Science *300*, 764.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. Science *326*, 1509–1512.

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.

Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960–964.

Capecchi, M.R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat. Rev. Genet. *6*, 507–512.

Carroll, D. (2012). A CRISPR approach to gene targeting. Mol. Ther. *20*, 1658–1660.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. *31*, 230–232.

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

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

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. *190*, 1390–1400.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

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. USA *109*, E2579–E2586.

Haber, J.E. (2000). Lucky breaks: analysis of recombination in Saccharomyces. Mutat. Res. *451*, 53–69.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell *139*, 945–956.

Horvath, P., Romero, D.A., Coûté-Monvoisin, A.-C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., and Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. J. Bacteriol. *190*, 1401–1412.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262–1278.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. *31*, 227–229.

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.

Jasin, M., and Rothstein, R. (2013). Repair of strand breaks by homologous recombination. Cold Spring Harb. Perspect. Biol. *5*, a012740.

Jiang, W., and Marraffini, L.A. (2015). CRISPR-Cas: New tools for genetic manipulations from bacterial immunity systems. Annu. Rev. Microbiol. *69*, 209–228.

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., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. eLife *2*, e00471.

Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., and 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.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. *13*, 722–736.

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.

Mangold, M., Siller, M., Roppenser, B., Vlaminckx, B.J.M., Penfound, T.A., Klein, R., Novak, R., Novick, R.P., and Charpentier, E. (2004). Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule. Mol. Microbiol. *53*, 1515–1527.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science *322*, 1843–1845.

Medawar, P. (1968). Lucky Jim (New York Review of Books), March 28, 1968.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. Nat. Biotechnol. *29*, 143–148.

Mojica, F.J.M., and Garrett, R.A. (2012). Discovery and Seminal Developments in the CRISPR Field. In CRISPR-Cas Systems, R. Barrangou and J. van der Oost, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 1–31.

Mojica, F.J.M., Juez, G., and Rodríguez-Valera, F. (1993). Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. Mol. Microbiol. *9*, 613–621.

Mojica, F.J.M., Ferrer, C., Juez, G., and Rodríguez-Valera, F. (1995). Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning. Mol. Microbiol. *17*, 85–93.

Mojica, F.J.M., Díez-Villaseñor, 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.

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

Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. Science *326*, 1501–1501.

Pandika, M. (2014) Jennifer Doudna, CRISPR Code Killer, Ozy.com, January 7, 2014. < http://www.ozy.com/rising-stars/jennifer-doudna-crispr-code-killer/4690 >.

Porteus, M.H., and Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. Science *300*, 763.

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.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. *32*, 347–355.

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.

Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R., et al. (2010). The primary transcriptome of the major human pathogen Helicobacter pylori. Nature *464*, 250–255.

Siksnys, V., Gasiunas, G., and Karvelis, T. (2012). RNA-directed DNA cleavage by the Cas9-crRNA complex from CRISPR3/Cas immune system of *Streptococcus thermophilus.* U.S. Provisional Patent Application 61/613,373, filed March 20, 2012; later published as US2015/0045546 (pending).

Sontheimer, E., and Marraffini, L. (2008). Target DNA interference with crRNA. U.S. Provisional Patent Application 61/009,317, filed September 23, 2008; later published as US2010/0076057 (abandoned).

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.

Sternberg, S.H., and Doudna, J.A. (2015). Expanding the Biologist's Toolkit with CRISPR-Cas9. Mol. Cell *58*, 568–574.

Travis, J. (2015). GENETIC ENGINEERING. Germline editing dominates DNA summit. Science *350*, 1299–1300.

Urnov, F.D., Miller, J.C., Lee, Y.-L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteusm, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature *435*, 646–651.

van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat. Rev. Microbiol. *12*, 479–492.

Wright, A.V., James, K., Nuñez, J.K., and Doudna, J.A. (2016). Biology and applications of CRISPR systems: Harnessing nature's toolbox for genome engineering. Cell *164*, this issue, 29–44.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell.

Zhang, F., Cong, Le, Lodato, S., Kosuri, S., Church, G.M., and Arlotta, P. (2011). Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat. Biotechnol. *29*, 149–153.

Zhang, F. (2012). Systems Methods and Compositions for Sequence Manipulation. U.S. Provisional Patent Application 61/736,527, filed December 12, 2012; later published as US008697359B1 (awarded).