Dorcas Cummings Lecture

Jennifer Doudna

Dr. Jennifer Doudna presented the Dorcas Cummings lecture entitled "Editing the Code of Life: The Future of Genome Editing" to friends and neighbors of Cold Spring Harbor Laboratory and Symposium participants on Saturday, June 1, 2019. Dr. Doudna holds the Li Ka Shing Chancellor's Chair in Biomedical and Health Sciences and is a Professor in the Departments of Molecular & Cell Biology and of Chemistry at the University of California, Berkeley, and the Executive Director of the Innovative Genomics Institute.

Thank you. It's a great pleasure to be here to give the Dorcas Cummings Lecture. I've heard this lecture a number of times when I've come to this symposium and I'm pinching myself that I'm actually up here doing this. I do have to tell you something important about Cold Spring Harbor. I was having dinner last night with a couple of colleagues and we were chitchatting about different experiences we'd had here at the lab. We realized that each of us at different times had given our very first scientific lecture here at Cold Spring Harbor, and for each of us it had been a transformative experience. So, I just want to point out that the lab plays a critical role not only in all of the incredible science that you've heard, but also for the next generation and helping people from all around the world come together and think and do science in a way that is really quite unique.

In that spirit, I want to tell you about an area of science that grew out of very small science from just a few labs around the world that were investigating an aspect of biology that initially seemed quite obscure: namely, how bacteria fight viral infection. When a bacterial cell is infected by viruses, these viruses inject their genetic material—their DNA—into the cell. Bacteria are rapidly dividing, rapidly growing organisms, so they don't have much time to defend themselves from this type of infection. A lot of the work in molecular biology over the last several decades has centered around understanding this warfare that goes on between microbes and their invaders.

One system for defense had been missed all those years, and it really came to light only when a few bioinformaticians and scientists in the food industry were studying microbes that are cultivated in making yogurt and cheese and things like that. They figured out that about 40% of bacteria and about 90% of other single-celled organisms called archaea have an adaptive immune system encoded in their DNA. These organisms can capture bits of viral DNA during an infection and integrate a little sequence of that DNA into their own genome at a place called the CRISPR [clustered regularly interspaced short palindromic repeats]. The CRISPR sequence has a distinctive property: It has a series of short repeated elements that flank unique sequences, and each of these unique sequences is an inserted sequence that comes from a virus. This is, effectively, a genetic vaccination card: a way that cells can record over time the viruses they've been exposed to by storing little snippets of the viral DNA in their genome. Importantly, next door to this locus are CRISPR-associated, or *cas*, genes. These genes encode proteins that, together with the CRISPR locus, constitute this adaptive immune system.

Here's how it works: The cell will make a transcript, an RNA copy of the entire CRISPR locus, a piece of RNA that is an exact replica of all the encoded sequences [pre-crRNA]. Those RNA molecules are processed into individual units that each include a sequence that comes from a virus as crRNAs [CRISPR RNAs]. These then combine with Cas proteins to form RNA-guided surveillance complexes that can search the cell looking for matching sequences of DNA. When those matches are found, they recruit the Cas proteins to cleave those targeted sequences and destroy them. If this is a viral DNA, it quickly gets made into mincemeat by this system. It provides a wonderful way for microbes to adapt environmentally to the infectious agents that they encounter.

This video illustrates this in a little more dynamic detail. This shows some bacterial cells that are being invaded by viruses; here's the infection occurring as the virus injects its DNA. If this cell has a CRISPR sequence in the genome, it can integrate a piece of that viral DNA into the CRISPR locus in this very distinctive pattern, flanked by these repeated elements; this is where the viral DNA sequences are stored. Then the cell is able to make an RNA copy, a transient molecule that is chopped into units, each including one of the viral sequences; these become the "zip codes" for the system. These RNAs then combine with a second RNA called "tracr" [*trans*-activating small RNA] and a protein called Cas9. That is the unit that goes searching the cell, looking for DNA that might have a match to the guide RNA from the viral insert. If a match is found, then the DNA unwinds and the Cas protein is

Published by Cold Spring Harbor Laboratory Press; doi:10.1101/sqb.2019.84.040139

able to cut the DNA at a precise position that, in bacteria, triggers destruction of those pieces of DNA. It's a great way to protect the cell from viral infection.

Initially, scientists were studying these and doing a lot of sequencing of bacterial genomes. What became clear was that there is not just one type of CRISPR system but, in fact, quite a few. This figure from a review from a few years ago [Fig. 2A in Mohanraju et al. (2016), *Science* **353:** aad5147] shows that we can divide these CRISPR systems into two different classes called Class 1 and Class 2 that are distinguished by the numbers of Cas proteins that are part of these systems. All of the Class 1 CRISPR systems have several proteins that are required for these systems to function, whereas the Class 2 systems in each case have one single large protein that is the only protein necessary for RNA-guided protection of the cell, and that was something shown genetically.

How did I get involved in all of this? I've been interested in RNA since I was a graduate student and I've been fascinated by how RNA molecules help cells to control the flow of genetic information, whether it's a cellular gene or a viral gene. When I heard from my colleague Jill Banfield about 12 years ago about the existence of these possible adaptive immune systems that might be RNA-based in bacteria, I was intrigued. We started investigating these and that eventually led me to a conference in 2011 where I met Emmanuelle Charpentier. Emmanuelle was studying a bacterial system that had this type of CRISPR element—the Type II CRISPR in Class 2—in the genome that included this gene called *cas9*. When we met up at this conference, we decided to team up to figure out the function of this protein.

It seemed like it must be an extremely interesting protein that would be RNA-guided and somehow targeting DNA, but at the time nobody knew how it worked so that was the question that we asked. What emerged from experiments done by our lab members, Martin Jinek and Kris Chylinski [Jinek et al. (2012), *Science* **337:** 816] is that Cas9 is an RNA-guided protein that uses the CRISPR RNA together with the tracrRNA to recognize double-stranded DNA at a 20-letter sequence that matches the sequence of the guide RNA. When that match occurs, it cuts *both* strands of the DNA double helix and makes a clean break, just like cutting a rope.

Martin Jinek in my lab was investigating this, and being a very good biochemist he was trimming away at these RNAs trying to figure out what was essential for this protein to function. Eventually, he realized that he could link together the CRISPR and tracrRNAs into what we called a "single-guide RNA" that would have both the targeting sequence and the Cas9 binding information in the very same RNA molecule. When Martin did this experiment and showed that you could easily program Cas9 with this single-guide RNA and direct it to cut a desired DNA sequence, we realized that this project had gone from being a curiosity-driven investigation of a bacterial immune system in our lab, which is where it started, to opening the door to something that could be a very important and exciting technology.

The reason for this is because of all of the other beautiful work that had gone on in many labs in the preceding two decades, showing that in eukaryotic cells—like ourselves, or plants, or animals, etc.—when there is a double-stranded DNA break that occurs in the genome, cells can repair those breaks. Rather than degrading DNA like what happens in rapidly growing microbes, in ourselves, those breaks are detected and fixed. Because of that, if you introduce a double-stranded break into a eukaryotic cell genome, you can trigger DNA repair that results either in a gene disruption or the insertion of new genetic information—of donor DNA—at exactly the site of the break.

Scientists had realized that if you could figure out how to introduce a double-stranded DNA break at a desired position you could trigger targeted genome editing, or engineering, as we called it at the time. There were a number of technologies that had been developed to do this—to introduce double-stranded breaks—that were exciting enough that a lot of people were talking about the possibilities of genome editing, but the technologies were too cumbersome for most labs to be able to adopt. CRISPR–Cas9, being a molecule that can be reprogrammed fairly trivially by changing its RNA guide sequence, emerged as a tool that could achieve this goal of cutting DNA easily in different kinds of cells to trigger genome editing.

This illustrates how we imagine that this works in a eukaryotic cell, where the DNA is inside the nucleus. Here's Cas9 with its guide RNA searching through the cell. We imagine that it is quickly binding and releasing DNA, looking for a match to its guide RNA. When that match is eventually found it latches onto the DNA and, by a mechanism we don't fully understand, it triggers melting and opening of that DNA helix to allow an RNA–DNA hybrid to form in the protein. Cutting of the two strands occurs and then repair enzymes in the cell come around, detect the broken DNA, and are able to fix the break by introducing a change to the sequence: in this example, by actually introducing a very small change in the DNA letters at that exact position. This tool has become a very powerful way for scientists working on any area of biology to make modifications to a desired genome to study fundamental questions, as well as to do things that are applied.

What happened next was truly remarkable. This figure from a journal publisher's website [https://www.elsevier.com/research-intelligence/campaigns/crispr] shows that in these years shown here, there were several technologies that had come along for genome editing—meganuclease, ZFN [zinc-finger nuclease], TALEN [transcription activator-like effector nuclease]—and there were some papers published using these technologies that were starting to populate the scientific literature. When CRISPR emerged, the publications using it just took off, and it's still going up exponentially. The reason is that this technology came along at a perfect moment when there was a lot of understanding about genomes, many sequenced genomes, and lots of exciting questions that scientists wanted to be able to ask about gene function. What was missing was an easy

way to manipulate those genomes, and that's really what CRISPR offers.

I want to turn now to the opportunities as well as some of the important challenges that are coming forward, given that we now have a tool for genome editing that is widely available, fairly simple and inexpensive to use, and quite effective for the kinds of applications that people are using it for. To start, as a scientist it's been incredibly exciting to see all of the ways that genome editing with CRISPR is being used to do fundamental research. There are many examples, but I want to point to one that was highlighted in the *New York Times*. If you were reading the science section last week, you might've seen a weirdly titled article about CRISPR snail babies. I looked up this paper and it's actually quite fun.

This is the picture that was on the cover of the journal *Development* [May 2019, **146** (9)], and it's all about these CRISPR baby snails. If you look carefully, these two snails look like mirror images of each other. Virtually all snails, for reasons that were unclear, develop with this right-handed coil to their shell and it was extremely rare to have animals occur in the wild that had a left-handed coil to their shell. Developmental biologists have been very interested in this question, partly because there are all sorts of examples across biology—including in humans—where there is a handedness to an organism and the genetics of that have been really unclear. These scientists, [Masanori] Abe and [Reiko] Kuroda, were able to use CRISPR first to look through the genome of the snail to try to figure out some of the possible genes that might be involved in this pathway. Then, using CRISPR, they could actually knock out or disrupt individual genes and ask what happens. They ended up finding a single gene that controls the direction of coiling of the snail shell. These are some images that they took of these snail embryos that have been either treated with CRISPR or not [Abe and Kuroda (2019) *Development* **146:** dev175976]. As the wild-type embryos develop, you can already see at an early stage that they start to coil in the right-handed direction. When they use CRISPR to knock out this single gene, they found that now these embryos were coiling in the left-handed direction, even at a very early stage of development.

This gives you a sense of the kind of question that scientists can now address using a tool like this, but I want to talk about three specific areas where I think we will likely see the impacts of CRISPR in our lives. These also raise important questions that we have to think about, both from an ethical and societal perspective as well as from a regulatory perspective. Of course, there are lots of commercial opportunities that have come along with genome editing as well that are opening the door for companies, but also raise challenges.

The first one is in the area of public health. Ever since I was a graduate student, people have talked about the possibility of "gene drives" in organisms. This is a nice graphic I found on the *Science News* website [https://www.sciencenews.org/blog/science-ticker/gene-drives-arent-ready-wild-report-concludes] that shows how a gene drive works. The left side shows you how normal Mendelian inheritance of genetic traits works, where you have a trait that's passed along to some progeny in a population, but not to all. If you have a gene drive present in this animal, however, this means that you can take a trait and hook it up to a genome editing tool such as CRISPR that's very efficient at introducing a gene into a genome. If you do that in the right way, this trait can be passed along and can propagate through all of the animals in the population very quickly, much faster than you would ever spread a trait with Mendelian genetics.

That sounds like a cool thing, but why do we care? Well, there's a reason that this is shown with mosquitoes as the example because there's an exciting possibility that you could use a gene drive catalyzed by CRISPR to spread a trait that would prevent it from being able to spread mosquito-borne disease in human populations. That would have an incredibly powerful positive impact on human health globally if it could be made to work well and safely. I Googled recently how much money has been invested in research in gene drives using CRISPR and just with three organizations—the Wellcome Trust, the Gates Foundation, and the Tata Group in India—$250 million has gone into this, and that's just what's been publicly announced. There's tremendous excitement about this, but also concern about the potential for this to get out of control if released into the environment. There's a lot of work and a lot of debate right now about how we encourage the science as well as ensure that it's deployed safely.

The second area—and this, in my opinion, is the area where CRISPR will have its biggest global impact in the near term—is in agriculture. Here's an example from a wonderful scientist right here at Cold Spring Harbor labs, Zach Lippmann, who has been using CRISPR–Cas in plants such as tomatoes to do very interesting things. This was a paper that he published a couple of years ago [Soyk et al. (2017), *Cell* **169:** P1142] in which he showed that you could use the CRISPR–Cas9 system in tomato plants to adjust one gene that controls the number of flowers and also the strength of the stems that hold the fruit on the plant. By doing this, you could actually control the yield of tomatoes without affecting anything else about them. Unlike traditional plant breeding, where random changes are introduced into seeds and then traits are selected—typically over months or years—with random changes that come along so we end up with roses that don't have thorns but also don't smell nice, here you have the potential to introduce a very desirable trait such as increasing crop yields but without affecting the quality of the fruit. This is certainly very interesting, and the genetics that are responsible for this in tomatoes apply to lots of other crops. You can imagine all sorts of other traits that you might want to be able to tweak or change in various plants to make them drought-resistant or pest-resistant or to make them more nutritious and things like that.

This is another example of work done at Penn State University. They were able to make a change to a single gene in mushrooms that prevents them from turning brown when you cut them open, very desirable potentially for commercial use. This work was the trigger for the U.S. Department of Agriculture to debate how they might reg-

ulate these sorts of products that might come soon to a supermarket or a farmer's market where we could all get access to them. Their conclusion was that if you made a knockout of a gene it would not be considered a genetically modified organism [GMO] because it does not include foreign DNA, and therefore would not be regulated in the U.S. [Waltz (2016) *Nature* **532:** 293]. That was interesting, but not all countries have come to that same conclusion. In Europe, the regulatory agencies decided that any plant that is manipulated by genome editing or any kind of genetic manipulation technology, even if you ended up with a genetically identical product at the end, would now be considered a GMO [https://www.nytimes.com/2018/07/27/science/gmo-europe-crops.html]. This has raised a lot of questions about just thinking fundamentally: How do we regulate this sort of thing? How do we explain to consumers or the public the difference between using a targeted approach to making a single genetic change versus introducing random changes as traditional plant breeding does? And how do we ensure that regulators understand enough themselves about the science so that they get it right and they don't limit the possibilities of this kind of technology in a way that I think will potentially be harmful to people? These questions are ongoing.

I want to turn now to things that are happening in biomedicine with gene editing. One of the interesting and really fun things about the CRISPR field is that as fundamental research on these enzymes has advanced it's been possible to understand enough about the way that these Cas proteins work to take advantage of properties that we didn't understand initially. Here's one example of research on a protein known as Cas12 that, like Cas9, is an RNA-guided double-stranded DNA-targeting enzyme. Cas12 has an additional property where, once it recognizes a target that's defined by its RNA guide, it turns on a *single*-stranded DNA cutting activity that can be useful for detection [East-Seletsky et al. (2016), *Nature* **538:** 270; Chen et al. (2018), *Science* **360:** 436]. This type of protein now can be used to detect the presence of DNA molecules that have a specific sequence by hooking it up with a fluorescently labeled molecule that gets cleaved only when the protein detects its target sequence. That could be extremely useful for what we call "point-of-care diagnostics": being able to tell somebody—maybe from a blood sample or other specimen in a doctor's office—whether they have a viral versus a bacterial infection, the details of what kind of infection they have, or even maybe someday things like if they have DNA circulating in their blood that contains a mutation associated with cancer. I think we'll have more of these coming up as people continue to investigate the fundamental biology of these systems.

I want to turn now to the primary way that people are imagining being able to use gene editing for clinical purposes. There are two kinds of genome editing that can be done. One is called somatic cell editing and the other is called germ cell editing. Somatic editing means making changes to the DNA in an individual that are not heritable. They cannot be passed on to future generations and they affect just that one individual. That's very different from what happens when we edit a germ cell, which would be a sperm or an egg or an embryo where the change to the DNA is heritable and it affects not only the individual, but also their offspring. This is very important. If you think about how genome editing will be used in the future—and even how it's being used right now—it gives humans the possibility of controlling the evolution of organisms or the way that they are studied in laboratories, including ourselves. As this has been contemplated, initially people focused on the opportunities in somatic cell gene editing, the nonheritable types. Going forward, at least for the foreseeable future, the vast majority of opportunities in clinical medicine with genome editing are going to fall in this category.

One example of how this type of technology is likely to be implemented in the not-too-distant future is sickle cell disease, a blood disorder that results from a single genetic change to the DNA. In fact, it's a single A–T base pair that is mutated to T–A in the hemoglobin gene in a blood cell of an affected individual. It leads to the production of a mutated protein with a valine amino acid instead of a glutamic acid that produces cells that have this very characteristic sickled shape. Because of this, these cells tend to occlude blood vessels and cause great pain to patients. We know how to diagnose sickle cell disease; we understand a lot of the biology of the disease, but clinicians don't have anything to offer to patients except for palliative care. Now imagine that you could use CRISPR genome editing to fix the disease-causing mutation. There are different ways that are already being used in laboratories to do this, and there are now several groups that are moving quickly toward applying to the Food and Drug Administration in the U.S. to do a clinical trial to test this in human patients.

There's a documentary called *Human Nature* that was made by Adam Bolt and Elliot Kirschner and kindly sponsored by the Simons Foundation [https://wondercollaborative.org/human-nature-documentary-film/]. The film begins with this boy, David, who is affected with sickle cell disease. He's a normal teenager except that every few weeks when he is affected by these crises, he has to be in the hospital. It's very moving and interesting and they use this as a way to tell the story of gene editing. I highly encourage you to see it when it becomes publicly available [now publicly available].

I think somatic cell editing is likely to be the most prevalent way that genome editing will be used in patients in the future, but there is also this potential to do heritable genome editing, meaning making changes that can be passed on to future generations. One of the very first papers that was published using CRISPR–Cas9 in animals was done in Rudy Jaenisch's lab showing that you could make modifications to the germline of mice very easily using this technology. This has now been done with lots of other kinds of animals—like pigs that are being engineered to be able to use their organs for organ donation, and lots of other examples. It was early in 2014 that a paper was published using CRISPR–Cas9 to edit monkey embryos that, for me, brought to the fore this idea of human germline editing. I wondered even at that time if

people were already moving ahead to do this because it seemed like it wouldn't be that difficult.

That motivated me to take a leaf out of the work that had been done by David Baltimore, Paul Berg, and their colleagues back in the 1970s to think about the ethics of molecular cloning. We organized a small meeting in California that resulted in this publication [Baltimore et al. (2015), *Science* **348:** 36] that argued for what we called a "prudent path forward" with genome editing. We did not use the word "moratorium" for a reason, but really this was effectively a call for a moratorium on any clinical use of human germline editing until the technology could be vetted and, more importantly, for the societal discussions to happen around the use of this technology in a way that would affect all of us, potentially, very profoundly. After this publication, there was at least one paper that was published that year in which CRISPR–Cas9 was used in a research setting in human embryos, initially in nonviable embryos but then eventually in viable human embryos, so it certainly showed that one could do this kind of thing. It motivated the National Academies as well as the Royal Society to get together and hold a summit in Washington and then eventually publish this report that came out in February of 2017 [National Academies of Sciences, Engineering, and Medicine. 2017. *Human Genome Editing: Science, Ethics, and Governance.* The National Academies Press, Washington, DC] that also argued for a moratorium on any clinical use of human genome editing.

In November of last year, I received an e-mail the day after Thanksgiving from this gentleman, He Jiankui, who informed me that he had in fact used CRISPR–Cas9 in human embryos, and those had been implanted and resulted in the birth of twin girls in China that had edited genomes. He presented his work at the international summit in Hong Kong on this topic, the second summit that was sponsored by the National Academies to discuss this. As you can imagine, this caused quite an uproar. One of the things that was very clear from his presentation was that it was relatively easy to use CRISPR–Cas9 in human embryos, as we had envisioned, but very difficult to do it well.

To illustrate this, I want to show you this figure [https://twitter.com/RyderLab/status/1068128997656207361/photo/1] from Sean Ryder at the University of Massachusetts. The top panel illustrates a gene called CCR5 [C-C chemokine receptor type 5] that encodes a protein essential for human immune cells to be infected by the HIV [human immunodeficiency virus]. That was the gene that Dr. He decided to target with CRISPR–Cas9 in these human embryos from parents in which the father was HIV-positive. It's been established that there's a few rare individuals in the human population that naturally have a deletion of 32 bp in this CCR5 gene that protects them from HIV infection [second panel]. The stated purpose of Dr. He's study was to introduce this mutation into these girls' genomes to protect them from future infection by HIV, except he showed these data right here [panels 3–5]; these were actual sequences from these girls' genomes. You can see that none of the changes that were introduced look like the naturally occurring deletion. So, yes, the genomes were apparently edited, but none of them received the change that occurred naturally in human populations. Instead, what was done here was to introduce changes that had never been seen in humans and never even been tested in animals, which really is quite shocking. It also emerged that the way that consent was obtained and provisions for following the health care of these girls after birth had really not been done properly either. On many levels, this came across to most people as deeply flawed and really wrong. This highlights why it's very dangerous to do this kind of genetic manipulation before we really understand our own genetics well enough to know how to predict the outcomes of these kinds of manipulations.

So, yes, there's a lot of talk about "CRISPR babies" in the media—babies engineered for things like high IQ, or 20/20 vision, or no baldness, or lower risk for a variety of diseases—but the reality is that we don't understand enough about the human genome and our own genetics to be able to do any of this. Most of these kinds of traits have not just one genetic cause; they have many genes that contribute to those traits. While CRISPR will help uncover those genetics over time, we're not there today. We also know that the technology is, frankly, not appropriate right now to be using in human embryos either. Fortunately, the World Health Organization and the National Academies of Science, Engineering, and Medicine have announced international forums that are going to be looking into this and making what I hope will be very restrictive guidelines that will be in place for anyone that might contemplate using CRISPR–Cas9 or any other genome editing technology in human embryos in the future. I think it's very important at this point not to call for yet another moratorium, but rather to keep a very open dialogue. There's a lot of interest in this and we can't put it back in the box; we really have to be open about it. People need to be discussing it actively, and I'm very glad that these international groups are going to look into it carefully. I hope they will proceed apace to do this in a timely fashion.

It's been incredibly interesting and exciting for me over the last few years to be involved in this field. I came to this from something very different. I'm an RNA biochemist and structural biologist. I was not doing anything with human genetics or any other kind of genetics. We were pulled in this direction because of the power of RNA-guided gene regulation in microbes, leading to the deployment of this technology across many other organisms. Both delivery and control of the DNA repair processes are going to be critical for this technology to go to the next level; we're working hard at Berkeley and with various collaborators on both of these aspects. Of course, fundamental research continues; we continue to understand more about the fundamental biology of CRISPR and other types of pathways and microbes. I personally feel very excited about uncovering what's out there in nature. It's an amazing time to be a biologist. In biology, we have fundamental questions that get asked in the laboratory and we uncover new information. Sometimes, that leads to new technologies, and those new technologies, in turn, allow yet more and new experiments to be per-

formed. Technologies and fundamental science really do go hand in hand.

I'm going to close by thanking a fantastic group in my laboratory: Yavuz Dagdas, Ahmet Yildiz, Eva Nogales, Jill Banfield, and Jamie Cate at the University of California, Berkeley; Emmanuelle Charpentier at the Max Planck Institute in Berlin; and Maria Da Costa and Joel Palefsky at the University of California, San Francisco. We're extremely grateful to the National Institutes of Health, the Paul G. Allen Frontiers Group, the Howard Hughes Medical Institute, the Defense Advanced Research Projects Agency, and the National Science Foundation for their financial support. I also want to call your attention to our Innovative Genomics Institute in the Bay Area. It's a partnership between Berkeley, UCSF, and the Gladstone Institutes that is fostering academic research, but we also encourage partnerships with companies to advance genome editing to solve practical problems, and you can check us out on our website [https://innovativegenomics.org/]. Thank you.