ADAM MARTIN: All right, so we're going to switch gears again today, and we're going to move off of kind of pure genetics and start to talk about molecular genetics. And I want to start with the concept of-- let's say you want to identify a piece of DNA, purify it, and propagate it so that you have it for future use. And so the process of doing this is known as cloning. And it's the process of, if you will, purifying and propagating a piece of DNA in an organism.

So sort of the goal for this lecture is for you to know how if you wanted to, let's say, identify a piece of DNA-- maybe you're interested in the piece of DNA that contains a given gene. How would you go about getting that DNA in a state that can be propagated sort of on and on? And how can you identify the piece of DNA you want?

And so one tool that we're going to use is we're going to use an organism bacteria as a tool. So I'll draw my sample bacteria cell here. Could be something like E. coli, just some bacterium. And you'll recall, when I talked about cells a few weeks ago, bacteria and prokaryotic cells have a chromosome called the nucleoid that's present inside. So that's the bacterial chromosome.

But bacteria can also have extra chromosomal pieces of DNA that are called plasmids that exist in their cytoplasm. These are plasmids. And these extra chromosomal pieces of DNA replicate independently of the chromosome. And they can persist in the bacterial cell and be passed on to the daughters of this bacterial cell as the bacterial cells divide.

So if we focus on an individual plasmid, what a plasmid would look like, often they have a cassette or a gene that confers antibiotic resistance. And that's often the reason that these bacteria are harboring these plasmids, because it gives them a sort of advantage if they're exposed to a certain antibiotic from a predator organism. So this would confer antibiotic resistance.

One common example is ampicillin resistance, which I'll abbreviate just amp with an R next to it. But these plasmids, for them to propagate from bacterial cell to bacterial cell, they also need to be able to undergo replication. So they also have what is known as an origin of replication, which is often abbreviated ori, which basically allows this plasmid to be replicated such that copies of the plasma are passed on to the subsequent generation of bacteria.

And we can take advantage of this plasmid system in bacteria, because we could take, let's

say, a piece of foreign DNA-- and this foreign DNA could be of eukaryotic origin. We could take a piece of eukaryotic DNA and insert it inside of this plasmid. And we can basically use the plasma as a sort of platform or a vector to carry the piece of DNA that we might be interested in and to use the bacteria to replicate that DNA and pass it on to its descendants so that you essentially have a clone of bacteria, and you have a clone of this DNA in a given bacterial population.

So, again, this would have an origin and maybe an ampicillin resistance to it. So how would you determine whether or not bacteria have a plasmid in it? Can you think of an experiment you could do to determine whether the bacteria has this plasmid? Stephen?

AUDIENCE: You could add ampicillin, and it'll survive with the plasmid.

ADAM MARTIN: Exactly. So what Stephen suggested is that if he wanted to know whether or not this bacteria had the plasmid in it, he would add ampicillin to the media. And if the bacteria doesn't have the plasmid, it won't be able to grow. But if it has the plasma, it encodes this gene that confers ampicillin resistance, and it will thus be able to grow. So that's exactly right. So you're able to select for bacteria with a given plasmid by simply growing it on media that contains an antibiotic.

I now want to go through steps in cloning a piece of DNA. And we'll go through it sort of as a series of ordered steps so you can see how the process works. I'm going to start with a step to cut the DNA. After cutting the DNA, we'll then mix pieces together. Once we mix the pieces together, we'll do something known as a ligation, and I'll explain that to you in just a minute. And then, finally, we'll end with selecting the plasmids that have the piece of foreign DNA that we want. And this is known as recombinant DNA, because you've recombined a piece of DNA from one organism-- it could be a eukaryotic organism, like humans-- with a piece of DNA that's from a prokaryotic cell, a bacterium.

So we then have some sort of selection process. So we're going to go through this step-by-step. And we're going to start with cutting DNA, OK? So, cutting DNA. And it turns out, we've talked about-- what type of enzyme do you think would cut DNA? Just generally, not as specific as what's up on the slide. What type of enzyme would cleave DNA? Think about how enzymes are named.

AUDIENCE: DNase?

ADAM MARTIN: Yeah, so Stephen suggested DNase, which is a really good suggestion, right? So an enzyme that will cut DNA would be a DNase. Another word for that is a nuclease. It's some type of nuclease. And the type of nuclease we're going to talk about is going to be an endonuclease. We talked about exonucleases, which chew DNA from the ends in the context of DNA replication.

But what an endonuclease is, is it's a nuclease that's going to recognize a fragment of DNA and cleave it in the middle. So it doesn't require an end. It's going to chop it right in the middle. And there's a type of endonuclease, and these are called restriction endonucleases. They are nucleases that our natively present in a lot of different bacteria. And these restriction endonucleases essentially look through the DNA sequence, and they recognize a specific sequence of nucleotides and make a cut right at that sequence.

So I have a few examples to show you here. The first is this EcoR1 restriction endonuclease. EcoR1's from E. coli, and it recognizes the sequence GAATTC. So it recognizes this six-nucleotide . Sequence and it cleaves the double-stranded DNA on the top strand between the G and the A and on the bottom strand between the G and the A, OK? And you can see that the two cuts are staggered. So when this cut is made, it leaves the DNA with two ends, and they're sticky, because there's a five-prime overhang at each end.

So each end has this TTAA sequence. And these nucleotide bases can base pair with the complementary sequence. So this sequence could base pair with a sequence that has an end AATT. So these two ends that I've generated here could stick to each other, or there could be other ends that have the TTAA sequence that could stick to them.

So another example is this Kpn1 endonuclease. And this is from a different bacterium. But again, it cleaves the DNA on the two strands. And this time, it cleaves the top strand farther down the sequence. And that generates what's known as a three-prime overhang. But again, you have an overhang. So this is what is known as having a sticky end. Because again, these nucleotides are available to base pair with a complementary sequence.

The final type of restriction enzyme that I'll tell you about is EcoR5, which is a different enzyme from E. coli. And this generates a break, but this time, it cuts at the same position on both DNA strands. And that generates an end that's known as a blunt end, because there's no overhang, and there are no nucleotides that would sort of basically recognize a complementary sort of end like the sticky ends do, OK?

So these are several of many, many different types of restriction endonucleases that are present in a wide range of bacteria. So then, now that you have a tool that allows you to cut DNA, you could then cut DNA from two different sources. And I've outlined that here. The vector is what the plasmid DNA is commonly referred to. So we commonly refer to this prokaryotic part of the plasmid, the vector DNA, and the part that we're trying to insert that's the foreign DNA, the insert. That's just kind of the lingo in the field.

So here, I have a plasmid. It looks like a plasmid. It has an origin of replication. It has ampicillin resistance. And it has this EcoR1 site, which just means that this DNA sequence has a GAATTC, OK? So it's something that will be recognized by this restriction endonuclease. And then if you cut this enzyme with EcoR1, you start with the linear piece, right? So I, at 6:00 AM, started engineering this here.

So let's say I have my plasmid DNA, and I cut it at the EcoR1 site. Then I cleave it. If you cleave a circle, now you have a linear piece of DNA, OK? But it has sticky ends, right? And these sticky ends-- so pretend that this is a foreign piece of DNA. This is my eukaryotic DNA. Let's pretend it carries the gene elastin. And this has ends, too. And if they're EcoR1 ends, then they will be able to stick to the sticky ends of the plasmid.

And if you just get one sticking, now you have this piece of DNA which is two different fragments in tandem. But it's going to be moving around in space in the cytoplasm. And at some point, it might be recognized and stick to the other EcoR1 end. And then you have now a circular piece of DNA again, but now your circular piece of DNA has this piece of foreign DNA that's present inside the vector, which is the sort of poster tape here.

So I just wanted to show you that. So you can kind of-- when you're doing molecular biology, you have to imagine the sort of end stick sticking to each other and how they're going to sort of wrap around and connect for the final product. OK, so let's say you get DNA, and your DNA could be eukaryotic DNA from, let's say, humans or flies or whatever your favorite eukaryote is. And in the genome of that organism, there will be many restriction sites. But if you chop it up, you will get various fragments that have sticky ends for EcoR1 on both sides.

And then if you mixed the vector and the insert together, you have some probability of getting that insert to be incorporated into the vector. And then once this is present and ligated together, you can then put this into bacteria, and it will be replicated. So we focused on cutting, but if you mix these together, like I showed you at the tape, you're going to have sticky ends

that come together and stick together.

And you'll eventually get a situation where you have your insert-- so you have your insert here that might have your gene of interest and your vector DNA here. But when these things initially stick together, you don't have a single molecule where everything is covalently attached, right? You just have these base pair interactions between the two overhangs as they stick to each other through base pair interactions.

So if we think about what's going on right here, you have, initially, if we're thinking about EcoR1, a sequence that is-- oh, sorry, this should be a C. This is the nucleotide sequence, but when it's sticking together, the top strand will have a free five-prime phosphate on this adenosine, and there will be no covalent bond between this adenosine and this guanosine. So that's what the top strand would look like. The bottom strand would look like this, where there are covalent bonds along the DNA backbone. Sorry. Had a mutation there.

And this is incorporated in a broader sequence. And this bottom strand will have a free five-prime phosphate here and a free three-prime hydroxyl here, but there's no covalent bond here. There's no covalent bond here, right? So, at this stage, you don't have a single piece of DNA. You have two pieces of DNA that are interacting with each other through base pair interactions. So, eventually, you want it to be a single piece of DNA. And so you have to perform a step that is known as-- sorry, my phosphate got in the way.

But you want to perform what's known as a ligation, where you take something that's just sticking together through nucleotide base pair interactions and you add a type of enzyme, which is called DNA ligase, to catalyze the formation of covalent bonds here and here. So DNA ligase is an enzyme that if you have a three five-prime phosphate here and a free three-prime hydroxyl, it'll catalyze the formation of a phosphodiester bond between these two bases and the DNA.

So this DNA ligase forms a phosphodiester bond. And you go from having no bond there to having a bond. So then you eventually would get this sequence now, where you have covalent bonds between each of the base pairs. And what you see here is you've recreated the EcoR1 site. So when you get two EcoR1 sticky ends sort of recognizing each other and sticking to each other and then you ligate them together, you recreate that nuclease site so that you can cleave it again with the EcoR1 enzyme.

So now, moving on. I'll move on right here. So the last step is that once we have pieces of

DNA with this insert-- and let's say we're trying to find a piece of DNA from a eukaryotic organism. We might start with an animal. We have to extract its chromosomal DNA, digest it with a restriction enzyme, and then we digest the vector with the same restriction enzyme. And then we're going to make-- we're going to randomly insert these pieces of DNA into different vectors such that each bacteria who gets one of these plasmids will be replicating a different piece of DNA that's of eukaryotic origin.

And this is what is known as a DNA library. So this is making a DNA library. And a DNA library is essentially a collection of different pieces of DNA that are from some source, OK? But different bacterial clones will be replicating a different piece of that DNA. So you can see the challenge now is to find the needle in the haystack, right? You're trying to find that one piece of DNA which is the one you're interested in.

And I'll talk about several strategies that you can use to find the piece of DNA that you're interested in. I'll focus on selection. But first, I just want to differentiate between two different types of ways you could search for a piece of DNA. You could do a screen. And this is similar to what we talked about on Monday, where you look through a whole population of individuals, and you look for a given phenotype.

So in the case of flies, we talked about how Morgan's lab was looking for differences in eye color. And that was a screen, because they looked through a ton of normal flies to find the one they want. Another strategy would be to do what's called a selection, where you kill off everything that's not what you want by making the organism grow in some sort of selective condition. And then you only allow the organisms to grow that are the ones that you want. So this is called a selection.

And I'm going to illustrate several examples of selections, just so that you get an idea of how this works. So the first example I'll give is antibiotic resistance. And as Stephen so kindly pointed out earlier in the lecture, the way we can select for bacteria that have taken up this plasmid is to select the bacteria that grow in the presence of the antibiotic. So let's say you had a population of bacteria and let's say this started out being sensitive to an antibiotic. You could transform them with DNA from a strain that's resistant.

And maybe that resistant strain has a plasmid that has a gene that confers antibiotic resistance, in which case, if it's taken up by this sensitive bacteria, if you then grow it on a plate that has the antibiotic, you might get a colony or a clone of cells that has taken up the

piece of DNA that you're interested in-- in this case, the piece of DNA that confers antibiotic resistance. Everyone see how that works? So you're selecting only the cells to grow here that have taken up this antibiotic resistance gene.

I'm going to use another example now from yeast, and it involves functional complementation. And I'm going to start with something that involves the biosynthesis of an essential amino acid. And then I'm going to go to a more interesting case, which is a case that involves an experiment that involved the identification of the master regulator of cell division in humans. But we'll start with just amino acid biosynthesis.

And there are mutants of yeast known as auxotrophs. And these are mutant yeast, or mutant microorganisms, that fail to produce an essential nutrient. So an auxotroph is a mutant that fails to make a nutrient that's essential. So fails to make a nutrient. And the opposite of an auxotroph is called a prototroph. A prototroph is essentially a normal-functioning microorganism that's able to produce all of the essential nutrients that it needs in order to grow and survive, OK? So this produces all nutrients.

And so let's say you had an auxotroph for leucine. So you had a strain that if you didn't provide leucine to it, it would fail to grow. So we'll start with a leucine auxotroph. And let's say you want to identify the gene that's defective in this leucine auxotroph.

You'd perform a similar strategy to this one, where you'd have your auxotroph that you're transforming, so your auxotroph down here. And you would transform that strain with DNA from what organism? If you're trying to identify the functional gene, what organism would you use to produce the DNA you're going to transform into that organism?

AUDIENCE: Prototroph?

ADAM MARTIN: Javier's exactly right. You'd use the prototroph, right? So in this case, you would use DNA from the prototroph, because the prototroph has a functional copy of that gene. You know it does, because it's able to grow without adding leucine. And then you could take the auxotroph mutants that's transformed with DNA from the prototroph, and you played it on media. And what should be a property of the media? Should leucine be present or absent? Carmen?

AUDIENCE: Absent.

ADAM MARTIN: It should be absent, exactly right. So you'd look on plates that lack leucine or minus for leucine. And you'd select for colonies that now are all of a sudden able to grow leucine. So

you've restored the function of leucine biosynthesis in this clone, and you've made it into a prototroph again.

OK, this is what's known as functional complementation, because you're taking a cell that is defective in some function, and you're complementing it. You're complementing or rescuing the phenotype, OK? Now, even as a former yeast geneticist, I don't find amino acid biosynthesis and functional complementation in the context of leucine all that exciting. So I want to present one last example that involves an experiment that is going to involve the yeast cell cycle mutants.

And I'm going to tell you about the experiment that led to the cloning of the master regulator of cell division in humans. And it involves a yeast mutant, and specifically, a yeast cell cycle mutant. And these yeast cell cycle mutants are what are known as conditional mutants. They are isolated as conditional mutants, meaning that these mutants are able to grow under certain conditions, but not others. And specifically, the condition they used is temperature, so they're temperature-sensitive mutants.

The yeast cells can grow at 25 degrees, but not at 37 degrees. So this is known as a temperature-sensitive mutant, where you can propagate the mutant at one temperature, but then you can see if you raise the temperature, then it stops growing. And so you can see the mutant phenotype, because normal wild type functional yeast can grow at both temperatures. So this is a special type of mutant.

And I'm going to tell you about an experiment done by Paul Nurse, who's an excellent yeast geneticist. And what he did was he used these yeast cell cycle mutants to identify the human gene for what's now known as cyclin-dependent kinase, or CDK for short. This is the master regulator of cell division in organisms ranging from yeast all the way up to humans, OK? But he used yeast as a model system to identify this gene.

And the process was he took yeast cells-- and Paul Nurse worked on fission yeast cells, which are rod-shaped cells. And he identified yeast mutants. Yeast mutants. And he had a mutant in the CDK gene of yeast. He didn't really know it at the time. But the yeast CDK mutant-- what he knew was that this mutant was critically involved in the cell cycle in numerous types of yeast. So he knew this is an important gene.

And what he wanted to do was to identify if humans had an equivalent gene that could function

in the same way. So if you just have this CDK mutant and do nothing to it, it will not grow at 37 degrees, OK? But what he did was to take a DNA library-- similar to what I showed you before, where you just chop up DNA from an organism. In this case, he's using a human DNA library. And he used a particular type of library, but I'm going to skip over that for now and come back to it later.

So he used a human DNA library. That's just a collection of pieces of DNA from a human source, OK? So he's taking human DNA, putting it into a yeast plasmid, and transforming yeast with that human DNA. And he's looking for a piece of DNA that's able to complement the CDK mutant, meaning the yeast cells would then be able to grow at the non-permissive temperature of 37 degrees.

So he then looks for, on a plate, colonies of yeast that are growing at the non-permissive temperature of 37 degrees. So if you didn't do anything with this mutant, if you didn't transform in the DNA, nothing would grow. But he identified pieces of human DNA that rescued the phenotype of this mutant, OK? And so these are yeast that have the human gene for CDK, and they now grow.

And this is a functional complementation experiment, because you're rescuing the growth of this yeast now not with a yeast gene, but with a human gene. And this human CDK gene is so conserved across eukaryotes that it's able to still function in a yeast cell, which is pretty amazing. So this just outlines the experiment here. At 25 degrees, these yeast mutants can grow and form colonies. And at that temperature, you can transform the yeast with different pieces of human DNA.

Most of the human DNA is not going to be what you want. You're looking for the needle in the haystack. So most of these are not going to grow at 37 degree. But you're looking for this guy here that gets the human CDK, and that restores growth to this mutant strain. So voila, you get a colony of cells that are growing. And boom, Paul Nurse wins a Nobel Prize and the rest of the yeast field, as well, or a number of people who are working on cell cycle mutants. This is one of the experiments that led to the 2001 Nobel Prize for a bunch of yeast cell cycle geneticists.

All right, so I've told you about how to find the needle in the haystack. And this was more common when we didn't know the genome sequence of an organism. But now I want to tell you how knowing the genome sequence of an organism would allow you to replicate and

amplify a piece of DNA in vitro. So I'm going to tell you about an approach known as Polymerase Chain Reaction, or PCR.

And what PCR is, is it's an in vitro method. So it's an in vitro approach to essentially amplify DNA. And so let's say you have a piece of DNA-- it could be a piece of DNA in the genome-- and you know the sequence of this DNA. And it has base pairs between the two strands. So, normally, for DNA replication to occur, what do you need? What needs to happen here? Can a polymerase get in now? No? Why not? Miles?

AUDIENCE: The DNA's going to-- because they'll try and [INAUDIBLE] away [INAUDIBLE] from each other, so you have to [INAUDIBLE].

ADAM MARTIN: Yeah, you have to unwind it, right? So you have to denature it first. So if you do nature it, now you have two single-stranded pieces of DNA, right? Now what would a polymerase need to replicate that? Yeah, Jeremy?

AUDIENCE: A prime.

ADAM MARTIN: A primer. Exactly, right? And if you know the sequence, you can have a company synthesize a primer that's the exact sequence here and base pairs here. And I'll just draw the five-prime end of the primer right there. And now this primer has a free three-prime hydroxyl here. And if you added a polymerase, it would synthesize this bottom strand here. So this is known as the forward primer.

And on the other strand, you can design a primer that's complementary to these bases here. Again, the five-primer end is out. This would be known as the reverse primer. And then you could have the DNA polymerase synthesize the opposite strand.

All right, so the step here will be to first denature. So the first step would be to melt or denature the DNA, double-stranded DNA. So you denature the double-stranded DNA. This is commonly done above 90 degrees Celsius. And then the next step is once you have these single-stranded pieces of DNA, you can act you can have a primer present that anneals to the opposite strands. So you can have primer annealing.

And this is commonly done between 50 and 60 degrees Celsius. You have to cool it down so that the primer can now base pair, such that not everything is denatured. So you have to cool it down for these primers to recognize their cognate sequence and base pair with it. And then once you have the primer annealed to the template, then you can add DNA polymerase to

synthesize a new strand. So DNA polymerize for new strand synthesis. And this is commonly done at around 70 degrees Celsius.

And then you can repeat this process over and over again. And at each step, you're going to double the amount of DNA that you have between these two primers. So let me just-- this is just a figure illustrating this. It's on the handout and online. Basically, you have your original double-stranded piece of DNA. You denature it and allow the primers to anneal. New strand synthesis.

Then you take these new pieces of double-stranded DNA, denature them. The primers anneal to those new strands, and now you get new strands. And you just keep doing this cycle over and over again, and you essentially amplify the piece of DNA that's between the two primer sequences. So this is often used in forensics, because you can have very little DNA, and just by adding primers, you can really amplify the number of pieces of DNA you have between these two fragments. So you go from having very little DNA to a lot of DNA.

OK, any questions about PCR? I'm going to move on to something that-- all right. I've really been focused on discovery up to this point. But I know that a number of you are engineers, and you probably want to engineer something. And so I've had to-- I'm going to tell you about a field that is moving so rapidly, I'm going to probably have to totally revamp my lecture for next year, OK? And I'm going to tell you about genome editing.

So the last part of this story, genome or DNA editing. And I'm going to tell you about a specific type of system called CRISPR-Cas9, which has been in the news a lot, and there's a lot of excitement about this approach in the context of editing the human genome and possibly curing genetic diseases. Who here has heard of CRISPR-Cas9? OK, good. That's good. Our media is doing its job.

So who knows what it is? OK, some of us know what it is. I just want to just give you a very superficial overview of what it is and why it's important. And I'm going to keep coming back to it during the course of the semester, because I think it raises a lot of ethical questions, and especially in the context of stem cells. I need you to know the foundation before we get into the really good stuff.

So, let's see. So we're going to engineer something. So we're going to talk about repairing DNA. And if we want to edit the genome, the way this is most often done is by making a

double-stranded break, OK? So if you make a double-stranded break in a piece of DNA, it can be repaired one of two ways. One is by non-homologous end joining, where the two pieces of DNA are basically just shoved back together again. And this results, often, in mutations.

So if you're trying to fix something, unless you're just trying to break it, that's probably not what you want. But an alternative approach to DNA repair that organisms have is something called homology-directed repair. In this case, you can break a piece of DNA and add a piece of DNA that has a different sequence, but with homology near where the double-stranded break is. And in that case, you can replace the original sequence with what you provide as donor DNA.

So it gives you an ability to essentially change the DNA sequence at a given locus if you're able to cleave the DNA at a specific locus. So, first, I want to start with just a thought experiment, right? You're all thinking, OK, we need to cleave double-stranded DNA. And boy, I just gave you a perfect tool for that. I gave you all these restriction endonucleases, right? So what's the problem with those?

Well, let's think about the human genome. The human genome is 3 billion base pairs. And an EcoR1 site looks like this-- GAATTC. So it's six nucleotides long. And if you think of just a random sequence of 3 billion base pairs, you would get this sequence randomly one out of every four to the sixth times. So that's going to be one every 4,096 times.

So if you get this in random DNA 1 every roughly 4,000 times, if you use it to cleave the human genome, it's going to cleave hundreds of thousands of places in the human genome. So we need much more specificity if we want to select, let's say, a given gene that has a disease-causing allele and try to fix it. Because if we use a restriction endonuclease, we just chop up the whole genome, and that would be bad.

So specificity is the name of the game here. This is not specific, and we need a tool that's more specific. And that tool is going to be CRISPR-Cas9. And what CRISPR-Cas9 is, is it's essentially an RNA-guided endonuclease. So it's RNA guided, and it's an endonuclease.

Restriction enzymes, right, they have nothing to do with RNA. They don't use RNA to recognize the nucleotide sequence. It's just a protein, and the protein recognizes the nucleotide sequence. In CRISPR-Cas9, you have an endonuclease, which is Cas9. Let's bump this up.

So the endonuclease is the-- the Cas9 is the protein. That's the endonuclease. But its

selection of a target depends on an RNA molecule that it's bound to, OK? So the specificity comes, at least in part, from what's known as a guide RNA, or single guide RNA. That's what's most often used in genome editing.

So this guide RNA basically allows this enzyme to find a specific sequence. And the guide RNA is 20 nucleotides, or looks for homology for a 20-nucleotide base pair sequence. So you can see, already, we're doing way better than the six base pair recognition motif. We have 20 nucleotides. And there are other components of the system which increase the specificity.

Then you have your Cas9 in blue, which is the endonuclease, your RNA, the guide RNA, in black, and the template here is in gray. And what you see is this RNA sort of exhibiting complementarity to this target sequence. And only if there's complementarity between the RNA and the target will this endonuclease get activated and cleave at this site. So the RNA is sort of serving like a guide dog for this endonuclease to guide it to a certain location to cleave.

So the idea, then, is if you want to edit the genome-- and why people are so excited about this these days is you now have a system that might allow you to generate a double-stranded break in one specific place in the genome. And if you can do that, then if you provide donor DNA that maybe has a different sequence-- if you consider a disease allele, right? Let's say you know there's a gene that when there's a certain allele causes an inherited form of a disease.

You could then take donor DNA from an unaffected individual and take cells from the affected individual and cut the locus that's problematic and get a repair of the defective allele using a normal allele of the gene. And that would essentially rescue the function of that gene if it were then reintroduced into the patient, OK? Do you see sort of roughly how this works?

So this is a very sort of broad and general sort of conceptual framework for how this might happen. Let's say you have an individual with a blood disorder-- let's say sickle cell anemia or beta thalassemia. Those are inherited blood disorders, which lead to anemia. You could remove cells, and what might be the best are the stem cells from a patient. And you could then take those stem cells and use CRISPR-Cas9 in vitro in cell culture to edit that individual's cells to repair the genetic defect.

And you could then reintroduce those to the patient, where if they're stem cells, they'd reoccupy the stem cell niche and produce functional blood cells that would then essentially cure the individual of the disease. This is how scientists are thinking about the use of the

system nowadays. And this hasn't really been successful yet, but there are several clinical trials that are currently underway, where people are trying to show that this can be used to treat human genetic diseases.

So in the next year, you are going to hear more about this, almost undoubtedly, as we start to hear the results of some of these patients. There are concerns about this, as well. I don't want to overblow it. There are certainly concerns. We don't know this is going to work. I mean, people have been talking about this type of stuff since I was a student 20 years ago. But I feel like we're getting-- we're much more advanced now, and the tools are more advanced. And so I feel like we're kind of getting to the point where there's a much greater chance that this will be successful these days than it was 20 years ago.

I just want to point out where this system-- how it was discovered and where it came from. And I like this as an example. Much like for the fly genes that defined major signaling pathways, this is a discovery that came from fundamental research on, basically, the ecology of bacteria. So this CRISPR-Cas9 system essentially evolved in bacteria as a form of an arms race between bacteria and their predators, bacteriophage, which are viruses that infect bacteria.

So this is an arms race between bacteria and their vicious predators, bacteriophage. And what CRISPR is, where these enzymes and this system evolved from, is this is a form of an adaptive immune system for bacteria, which is pretty wild in and of itself. So CRISPR is an adaptive immune system for bacteria. If you haven't gotten your flu shot, you should get it. We'll talk about human immunity later in the semester.

But this is where bacterial immunity kind of-- I'm sneaking it in. So the way this works in bacteria-- what CRISPR stands for is Clusters of Regularly Interspaced Short Palindromic Repeats. So you can see already thank god they gave it an acronym. Otherwise, it wouldn't be getting nearly as much buzz, because no one can say that.

And so where this CRISPR came from is on bacterial chromosomes of many bacteria, there's these clusters of interspaced short palindromic repeats, and the repeats are interrupted by spacers. And what researchers discovered are these spacers have sequence similarity and identity to sequences that are from bacteriophage. So each of these colored sequences here has some type of complementarity to some type of bacteriophage.

And so when a phage infects bacteria, or some bacteria, what happens is that there's a system to recognize that foreign genetic element and take a piece of it and insert it in the

genome. And that serves as a memory for the bacteria to remember that it got infected by that particular phage. And then, later on, what the bacteria does is it transcribes this region and forms these mature what are known as CRISPR RNAs, where you can see there's some sequence would recognize a foreign genetic element.

So, therefore, in the future, if this phage came around again, what would happen is one of these CRISPR RNAs would recognize the foreign genetic element through base pair complementarity, and it would know to cut it. And after the target is cut, it's then degraded by the bacterial cell. So this is a way for bacteria to remember what viruses have infected them and to have a defense mechanism against it. So it's a pretty cool system.

You know, what's also cool about this system is it's an adaptive immune system, similar to how we sort of recognize foreign pathogens. What's different about it is this is heritable. It's incorporated into the genome. And the more phage the descendants of this bacteria see, the more of these repeats you see. So this is a heritable immune system, which, unfortunately, we don't have.

So you should still get your flu shot. We'll talk about vaccination later on. Have a good few days, and I will see you on Friday.