MITOCW | 7.013-S05-L13

We're going to talk today about recombinant DNA. This is a new section of the class. We'll spend three lectures talking about recombinant DNA, just sort of a practical how-to guide with respect to manipulating DNA now that you've learned a bit about it and how we use it in research and in other aspects of medicine and agriculture and so on. Now, recombinant DNA has been around for a long time. And, frankly, it's got a bit of a bad rap. When people in the general public hear the term recombinant DNA they sort of conjure up these ideas of mad scientists going around, you know, cloning bizarre organisms or creating dangerous forms of life. Although that's possible, it's not what we typically do with this technology. In fact, it has some rather benign and other quite important applications. So, Claudette, I'm not going to get this board, am I? OK. Basically what we're going to be talking about is genetic engineering. That ought to make the budding engineers in the crowd happy. We're going to manipulate the DNA of organisms and actually transfer DNA from one organism to another for various purposes. There are three general tenants that you need to pay attention to here. The basic goal is to isolate, from a complex genome, a piece of DNA of interest. And amplify that piece of DNA, which could be a specific gene that you're interested in, or it might be some other sequence within that DNA. It doesn't have to necessarily be a coding gene that you're interested in. The purposes of this are to, for example, produce therapeutic proteins or compounds. In the context of disease, for example, where individuals have a genetic disease in which they cannot make a particular enzyme, you might be able to treat that disease by producing the relevant enzyme via this technology in another organism and then supply it to that individual. The organism that you put it in could be a bacterium that produces it in large quantities. And you purify the enzyme and then perhaps inject it into the individual. It doesn't have to be bacteria. You can actually use recombinant DNA technology, genetic engineering to produce a human protein in, for example, some farm animal. And there are actually companies that will introduce human genes into sheep or cows such that you can purify the recombinant protein from the milk of that organism. And perhaps it might be more folded more properly or modified more properly than it would be if in bacteria. And it might be more easily purified from the milk than it would be from the bacterial cells. We can also engineer organisms. So this is to make different life forms that have different purposes, different values to us. And I've given you the example in past lectures about genetically modified foods. And your book talks about this as well. You can make crops that are resistant to pesticides. You can make crops that are resistant to herbicides. You can make crops that stay fresher longer, taste better, have better shape for packing, all sorts of things by manipulating the genome of those organisms such that they now have new and useful properties. We'll also teach you about how to make transgenic animals. I've given you one example of this already where you might want to make a therapeutic protein in the milk of a transgenic cow, but there are lots of purposes for making transgenic animals. One important one, especially for me, is to make disease models. So once we understand a human genetic disease we can use that information to create animal species, most typically mice that carry those very same mutations. And then study the disease process in the context of a laboratory mouse in ways that are difficult or impossible to do in humans. And we'll teach you how to do that. It all flows from the same general principles of manipulating DNA and transferring it from one organism to another. The value of doing this is several-fold. Understanding our or other DNA sequences. So without this technology, we actually couldn't have done the genome project. We need to purify up a lot of our DNA in order to sequence it to determine its sequence. And likewise of other organisms. So this was instrumental in bringing us to the point we are today of understanding our genetic information to the nucleotide level. As I mentioned, this technology allows us to understand disease-causing mutations. Again, we've taught you about that mutations can occur, disabled enzymes or other proteins cause disease, but that doesn't tell you why. And using this technology to sort of dissect what's happening to that individual gene by manipulating it in different ways, putting it in different contexts we can begin to understand what the specific consequences of that mutation are, which ultimately allows us to deal with the disease more effectively. And another value of just manipulating DNA, which we'll come to in a future lecture, is what you hear about in the press as DNA fingerprinting. It's a very practical use of this technology in which you want to distinguish two individuals based on the specific changes in their DNA. And we can do that because we can amplify up DNA sequences from very, very small quantities and figure out the specific nucleotide sequences of those DNA fragments and be able to tell then two people apart. And this is useful in paternity cases, it's useful in criminal cases, it's useful for archeological purposes and otherwise. So there are really a great number of values to this set of technologies. So where do they come from? Well, we've been able to do this sort of thing since the late 1960s, early 1970s. And this is referred to, in our field, as the Recombinant DNA Revolution And there were three things that happened in this time period that made this technology possible. Firstly was the ability to separate genomic DNA, specifically into fragments. So genomic DNA, as you know, is huge. Our chromosomes are long, long linear pieces of DNA. Even bacterial chromosomes are a few million nucleotides in length. That's much too much to deal with for many of these applications. And so it was necessary to basically chop the genomic DNA up into manageable sized pieces and then be able to deal with those pieces individually. The second thing that was necessary -- -- was the ability to transfer the isolated pieces -- -- into another organism. And the organism of choice then and largely now is bacteria. So it was observations made in the 1960s that bacteria would actually take up from their environment pieces of DNA and sometimes incorporate them within their cells. That observation led to the notion that you could do that on purpose, not just at random, and that would be then using the bacteria basically as a cargo ship for the introduced DNA. And, finally, the ability to amplify the fragments of interest to large quantities. So if I take one piece of DNA from this guy, it's a single molecule, that's not enough. I cannot use that for much. I cannot use it to sequence his DNA, for example. I cannot use it to study specific properties of that gene. Instead I need to amplify it up to very large quantities. And, therefore, it was necessary to figure out ways basically to turn the bacteria into little DNA factories that would churn out large quantities of DNA of interest. And, again, we'll review how that was possible. And all of these things came together more or less the

of interest. And, again, we'll review how that was possible. And all of these things came together more or less the same time of this timeframe, and then the field took off. Now we could manipulate DNA very precisely and begin to understand DNA at a very specific level. So I'm going to give you some examples of what we do. I'm actually going to give you perhaps the first and only practical demonstration of this. We're going to do an experiment before you, a dangerous thing to do in front of a live audience. We're actually going to transfer a gene from one organism to another. We're going to start it today and we'll finish it in subsequent lectures. The gene of interest is a toxic gene. It's a toxic gene that's present in an organism known as S. pyogenes. You might know of this organism, might have heard of this organism because it's the flesh eating bacterium. Anybody heard of the flesh eating bacterium? This is true. If you get infected with S. pyogenes and you're slightly immunocompromised, but even if you have a bad strain of it even if you're very healthy, the bacterium will go about eating your flesh. It actually can be lethal it's so devastating. The organism, S. pyogenes is a bacterium that has a genome, of course. That genome has about a thousand genes. And the sequence of this genome is known. So by standard DNA sequencing technologies, we'll actually review those for you in a couple of lectures, specific DNA sequences of this organism is known. OK? So we know exactly which genes there are there. And, just for the sake of discussion, let's consider a couple of genes. Gene A, gene B and gene T. And it's this gene that we're interested in. It's the gene that encodes a very critical toxin that the bacterium makes that kills your cells. And so when you're badly infected by this bacterium, it's dumping out this toxin and it's killing the cells in the surrounding. And that's why it's flesh eating. OK? Now, you also heard in the previous couple of lectures ago that these genomes have other important properties. And just to remind you they have origins of replication, ORIs. All genomes, for the purposes of replication, all pieces of DNA that are going to be replicated need at least one origin of replication. So this genome has one. And this will come up in another context shortly, but what we're really interested in is this gene T here, the toxin gene. We want to clone it. We want to move it from this organism to another organism. And specifically the other organism of choice, and this is often the case, is E. coli. E. coli is a common gut bacterium present in all of you in very, very large quantities. It has been modified for laboratory purposes largely weakened so that it cannot easily make its way back into you and cause possible harm. It's used in laboratories in another context all the time. I'm depicting here an E. coli cell. This is always confusing because I depicted here with a very similar circle the genome of S. pyogenes. This is a bacterial cell. This is the membrane of the bacterium. And, of course, E. coli has its own genome which has roughly the same size and roughly the same numbers of genes. What we're going to do is transfer one more gene from this organism into that organism. Really, the purpose of today's lecture is how do you do that? But before we consider how you do it, let's ask the question why you would want to do it. Anybody have any ideas? Why would you want to transfer this potentially deadly toxin producing gene to another organism? Yeah? It could be that in a different context it's actually beneficial in one organism. You may be able to make a better E. coli for some purpose if you transferred this gene. My favorite idea, actually, is global terror. We can laugh about this but, you know, there are terrorists around who are thinking these very thoughts. Sad but true, people are making modified organisms with dangerous genes to make them more Ã¬pathogenicÃ® for deliberate release into populations. So I'm not teaching you how to do this for malicious purposes but just to know that there are people who do that sort of thing. Related to this point over here, we want to understand these genes for the purposes of biological research. We want to understand as much as we can about disease-related properties of genes and so on. And sometimes you cannot do the experiments you want to do in the organism in which the disease occurs. And it's easier to do it, in a sense, in isolation in another organism. And then another example might be to produce a vaccine. If I could produce this protein that's so toxic in large quantities, I might be able to make a version of it which is not toxic, very similar but not toxic. And then I could inject it into people, they would raise an antibody response against it such that if they were ever challenged with the real dangerous guy they would be immune. So that would be a very useful thing to do. And, again, if we were worried about the terrorists we might do exactly that. So this is the goal of our experiment. And we're going to start the experiment today. So what I've done is acquired from one of my colleagues in the Biology Department two tubes in which we have placed these bacteria. S. pyogenes is in this one. E. coli is in this one. There are bacterial cells in here that grow in suspension. They grow inside the liquid medium. If we want to isolate the DNA we first have to isolate the bacteria. So we spin these tubes in a centrifuge. The bacteria then pellet to the bottom of the tube. So the bacteria are now collected at the bottom of these tubes. We then decant off the liquid. I meant to bring a beaker, but I don't think they'll mind if I use this one. So it's always a little bit dangerous to do this, and we'll have to take care of this later because it is a bit dangerous. It's always a little bit dangerous to do this because you don't want to bump yourself like I just did. It actually is starting to hurt a little bit. So we'll get rid of the liquid here. And then we'll add various chemical solutions to isolate the DNA from this bacterium. Wow. It's really starting to go. I'm just kidding. Gets them every time. There really was no bacterium in there of any kind. I know. OK. But we're still going to go through this example. So we want to transfer this gene into this organism. What are we going to do? Well, the first thing we need to worry about is the fact that this gene, the T gene here, is contained on a very large piece of DNA. As I said, it's got a thousand other genes. It's about 4 million base pairs long. And I'm only interested in this one, so I need a way to isolate the T gene away from the other genes present in this genome. So for this purpose I need a tool. And the tool that I need is a restriction enzyme Restriction enzymes are nicely described in your book. They are enzymes present in bacterial cells which are designed to cut, to cleave DNA sequences. And they do so in a site-specific fashion. They are sequence-specific -- -- DNA cutting enzymes. We call them endonucleases, but you don't have to worry about that term. They are DNA cutting enzymes. Some of them produce, as you'll see in a moment, what we call sticky ends, based on exactly how they cut the DNA. And others produce what we call blunt ends because they cut the DNA in a slightly different way. There are hundreds of restriction enzymes that have been purified from various bacteria. There are now companies that will sell you the purified restriction endonuclease. There are whole businesses that are geared around selling this stuff to molecular biologists around the world. So you can order it up like you would order up a chemical from a chemical supply house. As I said, these are site-specific, sequence-specific DNA cutting enzymes. So they recognize particular sequences in the

these are site-specific sequence-specific DNA-cutting enzymes. So they recognize particular sequences in the DNA. One such enzyme, which was one of the first discovered, an enzyme called EcoR1, which comes from E. coli, that's why it's EcoR1, R1 is probably restriction enzyme number one, recognizes a particular sequence reading from the 5 prime end of the DNA molecule towards the 3 prime end. It recognizes the sequence G-A-A-T-T-C. And then the other side of that, of course, is a 3 prime end. This is the polarity of DNA that you've been shown in many cases before. Now what is the reverse strand? What does the opposite, the complimentary strand of this sequence look like? Can anybody tell me? You want to yell it out in unison. Here is a G and -- That was very good. Notice anything about this? Yes. It reads the same ways forwards as backwards. G-A-A-T-T-C. G-A-A-T-T-C. What do we call that? It's a palindrome. Your book uses the word MOM, which is a fairly boring palindrome. It is reads the same way forwards as backwards. I always like "A man, a plan, a canal, Panama." That's a palindrome, too. It reads the same way forwards as backwards. Anyway, many, not all, but many restriction enzymes recognize palindromic sequences. They read the same way forwards as backwards. OK? And because of that they cleave the DNA on both strands at the same position. EcoR1, this enzyme that recognizes this sequence will cleave the DNA between this phosphodiester bond, between the phosphodiester bond between the G and the A. And likewise on this strand. And the consequence of that is to produce a break in the DNA, so you now have 5 prime G, 3 prime C-T-A-A. Breaks occur here and the strands get separated. They get pulled apart. The hydrogen bonds that are holding these base pairs together are not strong enough to hold the two molecules together, at least most of the time, so most of the time they pull apart. So you end up with a fragment that has an end that looks like this and another fragment that has an end that looks like this, A-A-T-C, here's the 3 prime end, G, and here's the 5 prime end. OK? So you snip it at either side and pull it apart. Oops. I did something wrong. I missed a T, yup, on both of them. Thank you. I don't think I've ever drawn those right. In all the years I've taught this course, I always make it. Not the same mistake necessarily but a mistake. OK, so they make a break. And that then cleaves the DNA into two pieces. Now, importantly, at low enough temperatures these ends can find each other. And we call them sticky because they are complimentary to one another. They can actually reform those base pairs. So you would get a molecule which has 5 prime G, three prime C-T-T-A-A. And then base paired to this T-T-A-A would be an A-A-T-T sequence, which was covalently bound on this side to the C. No. Yeah, covalently bound on this side to the C and non-covalently bound to this G. OK? So there are non-covalent bonds here and here and there are base pairs in the middle that are holding these strands together. We call this annealing. And, again, at low enough temperatures you can get sticky ends of restriction enzyme cleaved sites to come together this way. Now, importantly another enzyme which we could use, discovered more or less the same time, is an enzyme called DNA ligase, which you would have heard about in the replication lectures. And DNA ligase will tie to non-covalently linked pieces of DNA together by a covalent bond. So DNA ligase will reintroduce this phosphodiester bond and this phosphodiester bond to produce now a completely covalently-bound sequence. OK? So we can cut the DNA at sites we want to. If we change the temperature we can get these DNA molecules to come together in a non-covalent fashion. And then, if we add this enzyme DNA ligase, they will become covalently bond once again. OK? Key enzymes in our tool kit. As I said, there are other enzymes that will cut DNA in a blunt fashion. They will not produce these sticky ends. I won't show you a new example of that today, and it's not terribly important that you know about them, but you can imagine that an enzyme that cuts in the middle between the central positions of the recognition sequence will make a blunt end. They won't be sticky. And those aren't as useful to us because the sticky ends actually help us get two pieces of DNA together in order to promote the cloning step at the end. OK. So the S. pyogenes genome that we need to isolate our T gene from has a very large sequence. It's 4.1 times ten to the sixth base pairs long all the way around the circle. And I was wondering. -- Well, let me just remind you that EcoR1 recognizes a 6 base pair recognition site. So how many EcoR1 sites are there in this genome? What's the frequency of EcoR1 sites along a piece of DNA? Well, it has to be this particular sequence, G-A-A-T-C-C, reading in that direction. The frequency of any given nucleotide at any given site is one in four because there are four nucleotides. The frequency of six of a given sequence in a row is one over four to the sixth, which is one in 4,096. So every 4,000 base pairs or so, at random, you will find an EcoR1 site. OK? They won't be equally distributed. There will be some parts of the genome where there are many by chance, other parts where there are a few, but on average there will be about one every 4,000 base pairs. And since the genome is 4 million base pairs long that means that there's going to be about a thousand EcoR1 sites scattered around this genome. OK? And they might be close together in some places, further apart in other places, and so on around the gene. Now, the T gene is in a particular place within the genome. I'm going to draw it blue. Here's the T gene. And when deciding what restriction endonuclease, what restriction enzyme to use to cleave the T gene in an intact fashion, I'll consult the sequence, which I know, I'll choose an enzyme which I know won't cleave this gene, won't cut this gene. There are no EcoR1 sites within this gene. OK? So there might be one on this side and there might be one on this side, but I've chosen EcoR1 because I know there are no G-A-A-T-T-C sequences within the T gene. OK? So to give you a visual depiction of what we're talking about, more visual aids, I brought with me a very, very dangerous piece of rope, a flesh eating piece of rope. They never fall for it twice, do they? In which I've depicted the S. pyogenes genome as a piece of covalently bound rope here. And the EcoR1 sites are shown in yellow. OK? That's really what they are. They're just little tags basically that the specific enzymes recognize. They're not there for the purposes of the enzyme recognizing them. They're just there. And the enzyme can recognize them. OK? And here I've colored in blue the T gene. And you'll see that there are no EcoR1 sites within the T gene. So now if I take this purified piece of DNA that I recovered from S. pyogenes and I put in solution with the restriction enzyme EcoR1 in the right sort of buffer and so on, the restriction enzyme will do what it's able to do. It will go along and find those sites and cut them. OK? Everywhere it sees one it will cut one. You know what I really should do? I should figure out how to put the rope back together. That would be a trick. I don't know how to do that. I should work on that. That would be great. I should do that. All right. So we've cleaved our DNA molecule into fragments. OK? And the S. pyogenes genome that we drew up on the board will likewise be separated into fragments. It will be separated into about a thousand fragments because there are about a thousand sites. OK? I

fragments. It will be separated into about a thousand fragments because there are about a thousand sites. Now I now want to separate those fragments from one another. In order to clone the T gene, I need to separate the T gene away from all the other genes and all the other DNA sequences in this genome. So how do I do it? Well, I take advantage of the fact that the pieces of DNA that I've liberated are different lengths. You can see that there are some short ones, there are some long ones, and it's actually quite easy to separate DNA based on its size. The way you do that is to introduce the DNA solution that contains the fragments of all sizes into what we call an agarose gel. It's like a slab of Jell-O. At one end, as shown here, here's the slab of Jell-O in which little wells have been cut out, little indentations have been placed, we introduce the DNA solution, usually with a colored dye to let us know what we're doing, and then we put buffer on top of the gel and at the two wells at the end of the gel, and then we apply an electric field. This is called gel electrophorisis. We apply a positive charge on one end, a negative charge on the other end. Which way will the DNA go? Which way will the DNA go? Is DNA positively charges or negatively charged? It's negatively charged because all those phosphates which carry a negative charge. So in neutral pH buffer it will move towards the positive pole. And so the DNA will actually separate. And that's what you're seeing here. These are gels of different concentrations of the stuff that makes them stiff. So at a low percentage you see this fragmentation of fragments. The larger fragments don't go very far, the smaller fragments go further into the gel, and at different concentrations of this stuff you'll get different separations in different positions. And so with this separation technique I can isolate this fragment, which is of a particular size, away from other fragments. So if you imagine this is the gel here, I put a negative charge here, a positive charge here, the different DNA fragments are going to separate along the gel with the long ones at the top and the short ones at the bottom, and the middle sized ones will separate in the middle and the T gene will separate right there. And, if I have a lot of DNA to start with, I'm going to have a lot of molecules of the T gene in this position on the gel. And if I shine a UV light, which was what was done here, I can literally see that position. And since I know how big it is, because I know the sequence of this organism, I can know that it's exactly 1,512 base pairs long. And, therefore, it should run right there in this particular position on the gel. So I can purify it. Now, usually when I purify it, I purify a bunch of other things that are of similar size. So I don't usually get it all only this molecule, but I do a pretty good job of purifying it. And now I need to amplify it. I've got some of that DNA, but I need to have more of the DNA. I need to have much, much, many, many more copies in order to be able to do the manipulations I want to do downstream. So the next thing I need to do is get this piece of DNA, and the other ones that I'm less interested in, into bacteria and get the bacteria to make more and more of it. So what I do in principle is to introduce the DNA into different bacteria. There you go. And now I need those bacteria to make more of that stuff in order to give me what I need. But there are certain problems. There are certain problems associated with what I've just told to you. The problems are that I cannot tell which of these folks got the right gene. That's one problem. I need to figure out a way to know who got the T gene and who didn't. I also need to know who got any piece of DNA versus who got nothing. The rest of you didn't get any DNA. You are of no interest to me. If you were bacteria, I wish you would die. It's actually true. You'll see in a second. The rest of these guys I'm somewhat interested in, and I'm really interested in him because he's got the TG. I've picked on you a lot, haven't I? So those are two of the problems. And the third problem is that I need these guys to divide, which they're not terribly willing to do. I need them to divide, but bacteria will divide. But the bacteria, if these guys were bacteria, wouldn't know what to do with that little fragment of DNA that I gave them. They wouldn't know how to replicate it, how to make more if it because I've just given them a naked piece of DNA and that's not terribly useful. So we need to overcome that. And I'm going to tell you how. The process that we just described, the introduction of DNA into the bacterium is called transformation, bacterial transformation. Again, here's a bacterial cell with its own genome. We treat the bacterium with chemicals that sort of loosen up the membrane a little bit or we electroshock the bacteria which blows holes in the membrane. And then when there are holes in the membrane, these little fragments of DNA which I'm representing here, maybe I'll just show it as a single linear piece of DNA, can float in inside the cytoplasm of the bacteria. They literally just float in there. OK? Now, the bacterium, if they knew what to do with this piece of DNA, are remarkably good biofactories because bacteria will divide at optimal conditions three times every hour, three divisions every hour, which means 72 divisions every day. And if they could duplicate this DNA every time they divided, if they could and they were able to divide 72 times in a day, if there were one copy of that piece of DNA per cell at the beginning of the process, just one cell that had that piece of DNA in it then at the end of the day there could be two to the seventy-second molecules of DNA. OK? 72 divisions. Two to the seventy-second which is roughly equivalent to ten to the twentieth copies. Now, each of you has ten to the thirteenth cells in you. So you've got ten to the thirteenth or so copies of any particular piece of DNA. This is ten to the seventh times more than that, so that's 10 million persons' worth of a given piece of DNA. It's really quite remarkable. Now, Claudette always makes me point out of the fact that you never reach this theoretical limit because you would need an MIT's worth of bacterial solution to do it. So in practice, when we do this sort of experiment, we get about ten to the tenth to ten to the thirteenth bacteria at the end. And, if each bacteria were able to copy this DNA faithfully, we would have about ten to the tenth or ten to the thirteenth copies of the DNA. Still quite impressive for not a lot of work. The problem I've already alluded to is that the bacteria don't know how to deal with this introduced piece of DNA. This piece of DNA, for it to be brought up to ten to the tenth or ten to the thirteenth copies needs to be replicated. The DNA needs to be replicated. What do we know about replication? What did you learn about replication? What is the key thing that a piece of DNA needs in order for it to be replicated? An origin of replication, right? And these fragments most likely are not carry an origin of replication, so they're not going to be replicated. So for this purpose we need an origin of replication. And you'll see how we accomplish that. We also mentioned the fact that most of the bacteria, when we do this transformation, this is actually a relatively inefficient process. So most bacteria get no DNA. And they are, of course, not interesting to us. That's all of you guys in the back there. So we need to get rid of all of you. And then we need to find which bacterium has the clone of interest. Which bacterium has the piece of DNA of interest? This guy compared to the rest of them. So the solutions to all of those problems come in the form of what we call vectors. The piece of DNA that we introduce at the beginning

of these problems come in the form of what we call vectors. The piece of DNA that we introduce at the beginning is not an isolated fragment of DNA, in fact. Instead, we introduce the piece of DNA in the context of another DNA molecule called a vector. Vectors, for the most part, are derived from naturally occurring small DNA molecules that are carried in many bacteria outside of the main bacterial genome. This is the main bacterial genome. Many bacteria carry within their cytoplasm small copies of DNA, also circular, called plasmids. And these are short. They might be on average 5,000 base pairs long compared to a couple of million base pairs long. So they're small circular pieces of DNA. They get transferred from bacterium to bacterium in the wild for the purposes of genetic exchange between bacteria, also for transmitting drug resistance or pathogenic properties from one bacteria to another. So they do get exchanged from bacteria to bacteria. They were discovered a long time ago and it was recognized that they would be very useful as vehicle, vectors for cloned DNA. They have three key features, which I will illustrate in this diagram. One is the origin of replication. So they do replicate inside bacterial cells. They have their own origin of replication. So you can see where we're going here. We're going to put our piece of DNA, along with this, in order to get the replication of our DNA as well. They have restriction sites. Of course, they're going to have restriction sites throughout their DNA sequence. But they're designed now on purpose to have particular restriction sites in a particular region of the piece of DNA which we use for cloning purposes. So there might be EcoR1 site here and nowhere else in this plasmid, another enzyme represented once, and so on and so forth. Because we're going to introduce our DNA into here and we don't want to disrupt the sequence of the rest of the plasmid. And the final thing they have of value is a drug resistance gene, sometimes multiple drug resistance genes. An example of a drug resistance gene is the amp resistance gene, and this confers resistance to a common antibiotic called ampicillin, very similar to penicillin. If you treat most bacteria with penicillin they will die, with ampicillin, sorry, they will die. If you treat a bacterium that's carrying a plasmid that has an ampicillin resistance gene the bacterium will live, OK, because it inactivates the ampicillin. It makes it non-toxic to the bacteria. And now you can see how we're going to distinguish bacteria that get DNA from bacteria that don't get any DNA. The bacteria that get DNA are going to get a drug resistance gene. So they're going to live in the presence of ampicillin. If you don't get any DNA and you get treated with ampicillin you're going to die. OK. So, quickly. I've taken my S. pyogenes genome and I cut it up into EcoR1 sites. I've cut it with EcoR1, so I've cut it into EcoR1 fragments. And now I'm representing them as double-stranded with their sticky ends showing. Here's the end of an EcoR1 site. Here's the end of an EcoR1 site. I'm just going to show you two for the sake of simplicity. Here's another fragment liberated by EcoR1 from the S. pyogenes genome. Maybe these are two of similar size and I purified them together when I used my agarose gel technique. So these are two pieces of DNA. And this one contains our friend the T gene. We've liberated these, they have their sticky ends, and then we mix them with a plasmid that likewise has been cleaved with EcoR1. So we have lots and lots of cut plasmid that has the complimentary sticky ends. OK? This plasmid is that thing. I'm showing it as double-stranded now so I can emphasize the sticky ends. When I lower the temperature, at some frequency this DNA molecule with find its way here, an anneal at this end over here, at this end over here, and this piece of DNA will find its way to a different plasmid an anneal over here and over here. If I wait a while for that annealing process to take place and then I add DNA ligase, the enzyme that seals the nicks that form when those things come together, I will get fully closed double-stranded DNA molecules which are the plasmid plus a new piece of DNA. This one will give some chunk of the S. pyogenes genome I don't care about. This one will give a plasmid that has a piece of the S. pyogenes genome that carries the T gene. This one I do care about. OK? So, in the final minute. I now take these plasmids that I've generated by this technique, these recombinant plasmids, and I introduce them into bacteria in culture using this transformation technique. And then I plate them onto a Petri dish. And the Petri dish, importantly, has agar for the bacterial colonies to grow and it contains ampicillin. It contains the drug ampicillin. Now, as I said, the transformation process is inefficient. Most bacteria that I treat in the transformation process will get no DNA. They'll land on this plate. And what will happen to them? They'll die. The plasmids that did get DNA, sorry, the bacterium that did pick up a plasmid will have picked up an ampicillin resistance gene. They will grow. And they grow on little colonies, into little colonies after about a day which have about a million cells each. And they're separated from one another. So this colony might have that plasmid in it. Every cell in this colony has that plasmid in it. And this colony might have that plasmid in it. OK? I'm interested in this one because I want to now purify large quantities of that piece of DNA. I'm actually not interested in this one. And next lecture we'll talk about how you specifically purify this, isolate and identify this colony compared to the rest.