ERIC LANDER: Good morning. Good morning. So we've been talking about recombinant DNA. And really what it does to our picture here-- function, gene, protein-- is for the first time take something that's a theoretical relationship and make it operational. Being able to go from a function like the ability to make your own arginine, to a specific gene, to a specific protein, and to be able to connect those up.

In principle, by the time we're done with recombinant DNA, one should be able to go from any vertex of that triangle to any other vertex of that triangle. Given a function, find the genes. Given a gene, find the proteins. Given a protein, find the genes. Given a protein, find the function.

That's really the goal of recombinant DNA is to be able to start at any vertex and reach any other vertex of that triangle. We're not there yet. But we will be in the next couple of days, to the point where we can move freely about this whole picture.

So we've talked about DNA sequencing. And I want to pick up a little bit with DNA sequencing. And I'll probably end with DNA sequencing again, as I tell you where things stand today.

How did we use DNA sequencing? Well, we found us a clone. Maybe our clone was a clone that conferred the ability to grow without arginine. That was cloning by complementation.

Maybe it was a clone that encoded beta globin. There we found an antibody that recognized beta globin. We made a cDNA library with an appropriate promoter. And we asked E. coli to produce those human proteins from those cDNAs. And then we use our antibody to recognize which clone.

One way or the other, we found ourselves a clone. The clone had this vector. It had this insert. The insert is now of interest to us. We wish to sequence it.

And we talked before about taking that piece of DNA and subjecting it to sequencing. We started with a primer. From that primer, we extended. And we hit a

point, let's say, opposite an A or maybe opposite the next A or opposite the A after that or opposite the A after that. And we had this clever trick, for which Fred Sanger actually won a Nobel Prize, of using a defective version of the nucleotide T that would stop at that point.

But remember, we didn't use only defective T. We used a mixture of good T and defective T. Let's say defective T was 1% of the whole mixture, whenever we encountered a defective T and put it in, the chain would stop because it couldn't be extended. Whenever we put in a good T, it would keep going.

And so since we're making-- Of course, we have millions and millions of copies of our template sitting there in the test tube. We're always working-- when I talk about "a" molecule, and I draw a picture of a molecule, we should always know that there's millions of copies of those things. Some of them are stopping here. Some of them are going here, et cetera, et cetera, et cetera.

And we end up with a large collection. And if we separate it on a gel, we can detect the lengths of those fragments. If we attach a fluorescent dye, then those fragments are fluorescently labeled. We could all run them in the same lane with different colors. And we could put our little fluorescence detector and see what goes by.

And the traces you would get, the actual pictures that emerge from this, look something like this. And you'd see colored traces like that emerging from that electrophoretic detector. And you could read off the sequence by the colors-- really very gorgeous, very simple, beautiful technology, all this taking place in a little capillary tube.

You remember how we made a defective T, right? How did we make a defective T? Sorry?

STUDENT: Dideoxy.

ERIC LANDER: To dideoxy. Because remember, we need that 3-prime hydroxyl in order to extend it. No 3-prime hydroxyl, no extension. So if you make that deoxy at the 3-prime position, you can extend.

And since it was originally 2-prime deoxy, it's now 2-prime, 3-prime dideoxy. That's all it takes in order to block the extension reaction. And they sell the stuff, and you can use it.

All right. I had another question for you guys. Where did the primer come from? Sorry? Sorry?

STUDENT: We put it in.

ERIC LANDER: We put it in. How did we know what to put in?

STUDENT: [INAUDIBLE] catalog.

ERIC LANDER: Well, the catalog's very smart, but it doesn't tell us what we need. How do we know what sequence it is? Actually, it turns out that so many different sequences might be needed for different purposes in molecular biology that they don't stock them all in the catalog because you might order any sequence.

So it really turns out that if you want to order a specific 20-letter sequence, you go on the web, type it in, and then the machine will make it for you, and you get it the next day, it turns out. But they don't actually put it in the catalog because it would be too big an inventory. Although ones that people use a lot they keep in the catalog, otherwise they just make it on the spot for you.

But how do we know what sequence we're supposed to use? Here's my clone, let's say arginine, the ARG1 gene. How did I know what sequence to start with? You let me get away with just put a primer there. And what does it match?

Yeah?

STUDENT: Could you use the EcoR1 [INAUDIBLE]?

ERIC LANDER: I did use EcoR1.

STUDENT: [INAUDIBLE]

ERIC LANDER: EcoR1 site here, GAATTC-- GAATTC-- I'm going to cut in this site. I'm sorry, I'm going to cut in this site like that, let's say. Yeah, like that. And this fragment here-- sorry, like that-- the fragment here that I'm going to start sequencing from it starts with a G, right? Because it's opposite that C.

So because I cut with EcoR1, I'm pretty sure it starts with a G. It's not a very big primer to use, though. I could start with a C, but I don't think that's going to have enough binding energy to do it.

STUDENT: Doesn't a bacteria already replicate that DNA?

ERIC LANDER: It replicates that DNA just fine.

STUDENT: So there's already a primer in there. Why do you [INAUDIBLE]?

ERIC LANDER: Well, because I've now purified out my insert. And I'm going to subject it to sequencing. And I need to start with a primer. I've got that fragment. I need to know how that fragment starts. And I don't know how that fragment starts.

STUDENT: Is there a place you could cut in a little bit after that?

ERIC LANDER: Oh! What if I was really smart and put in a different restriction site back here? And I use that restriction enzyme, and I cut there? Then what would you be able to tell me? Well, then the fragment would start with a known vector sequence.

Bingo. That'll work. Good. Good engineering.

Since I don't know what the sequence is of the thing I'm reading, I'd better back up a little bit and use a known sequence from the vector. And then I keep going. This is all just to give you a sense of all the tricks you can do.

So now it's easy. Now, in fact if that's the vector I use a lot, all the time, then bingo! I can go to the catalog because they will stock that one. And I'll use it. And I can use my green primer there to get going.

All right. Now here's the problem. As I start sequencing, coming down this capillary

tube are fragments of different lengths. The speed of migration depends on the logarithm of the length of the fragment.

Big fragments goes slower. It's inversely proportional to the log of the length of the fragment. So big fragments go slower. And as they get bigger and bigger and bigger, they go slower and slower and slower.

But the difference between log of 1,000 and log of 1,001 is pretty small. And then log of 1,001 and 1,002, very small. Actually, those peaks over there start bunching up, and I can't tell them apart. So I can actually only go with this electrophoretic process maybe 1,000 letters before the peaks get too bunched up because the different speeds are not so different anymore. So I can only read 1,000 bases, let's say.

In practice, we would tend to read 700, 800 bases because it started getting scruffy. But let's make it round, and we'll say 1,000 bases. Now, suppose this fragment that I got that's the ARG1 gene is 3,000 bases. Well, we've got our clever trick that you've introduced here of cutting back over here at some previous site, using our primer here and reading.

But it kind of dies at about 1,000 bases. I can't read any further. What do I do?

STUDENT: [INAUDIBLE]

ERIC LANDER: What if there's not a perfect restriction site there? But you're on the right track. Keep going. Once I've sequenced the first 1,000 bases, what do I know? The sequence of the first 1,000 bases.

What primer could I use then?

STUDENT: [INAUDIBLE].

ERIC LANDER: I can just make a new primer based on those bases, right? So I don't even need my restriction site anymore. You've got it exactly right. I use my knowledge.

And then I could use a new primer, and I could go more bases. Then I could use a

new primer and go more bases. Then a new primer, and go more bases. And I can do what's called "primer walking" along the clone.

Will that work? You bet. That works just fine.

It's also very slow. Because I had to get my first bases, analyze them, order a new primer, and the next day set up my next reaction. Take a couple days, get my next reaction. Take a couple days, next reaction.

Imagine sequencing the human genome like this. This could take a long time if I do it in serial. So what else could I do? This works, by the way.

This totally works. It's a good procedure, and it's used for certain purposes. But what else could I do?

Here's the cool thing. We've got biology. But you guys, being MIT students, you've also got computer science and other tricks available.

Here's a cool trick. I have my clone, 3,000 bases. I like my clone. I'm going to take my clone, my fragment, of 3,000 bases, and maybe instead of sequencing it, I'm going to shred it up into a lot of smaller pieces.

Suppose I shred this up into fragments of, say, size 800 bases. This was 3,000 bases. Now let me shred it up. I'll just take 800 as a number.

Now remember, I had a lot of copies of this clone, right? So I'm going to get shreds like this and shreds like that and a shred like that. It's not just one copy. I've got a lot of copies of this piece of DNA there. I'm going to shred it up.

And there are ways to shred it up by forcing it through a needle or treating it meanly, or things like that. I'm going to get lots of little fragments. And what I could do is I could clone all of those little subfragments.

I take my big fragment, and what I'm going to do is I'm going to make a new library of subfragments. Got it? Now I have a whole lot of subfragments, each taken from my 3,000 bases. And they're all kind of smallish.

What could I do to all of those little subfragments? They're all living in a vector. Each is in its own vector.

I've spread them out. I've made a library. They're each in their own bacteria. They're each in a vector. That vector has a known sequence at its end, the green primer.

Couldn't I just sequence a lot of different random subclones and paste them together? See, that's where it pays to have computer science as well. Because what I could do is by subcloning these into individual little random pieces-- I have no idea how they've been broken up. I don't care-- I just take the total DNA, shred it up into pieces, sub clone it into a vector.

And now because it's in the vector, I could read this one and that one and this one and that one and this one and that one. Do I actually know which ones I'm reading? No. It's totally random.

I take my 3,000 bases, shred it up into lots of smaller pieces, and I just read a lot of them. When I get a whole lot of these pieces, maybe 800 letters each, what do I do? I write me a piece of code that looks for overlaps between them and start pasting it together. And that's called assembly. You assemble the sequence out of its little pieces.

And so you can assemble things. And this gets referred to as shotgun sequencing. That's what it's really called, because it's like you shoot it out of the end of the shotgun or something. It's broken up into a lot of pieces. It's just a shotgun, random approach where I take individual random clones, and I assemble them.

Any questions about that? It's really a way to do it. And the big difference there is you can do it in parallel. Rather than doing it one step at a time, which sounds so logical but takes so much time-- easier. Just shred it up, read lots of them all the same afternoon, and then assemble them by computer. So that's nice.

So I do that. I get my clone. I'm going to now do my computer assembly of it. And

I'm going to get my 3,000 bases. How do I analyze my clone, my clone sequence? Now I have 3,000 letters in order, nicely done. What do I do with it?

That clone, let's say, is able to confer the ability to grow without arginine. It encodes some enzyme that lets you make arginine. I've got 3,000 letters. How do I tell what it's doing? What do I look for? Yep?

STUDENT: Compare it with something that doesn't have [INAUDIBLE]?

ERIC LANDER: Well, so tell me what I'm looking for? I'm looking for a gene?

STUDENT: Yes, [INAUDIBLE].

ERIC LANDER: Yeah. So what about that gene? What's distinctive about genes? How do I recognize a gene? It's tricky.

STUDENT: The sequence?

ERIC LANDER: How can I just recognize it from the sequence? Can I tell that something is a gene?

STUDENT: Start codon.

ERIC LANDER: I could look for a start codon, ATG. Do you think that'll happen just by chance, though? There'll be a lot ATGs running around, because you've got two strands, three reading frames. It'll happen pretty often. But that's a start.

What happens after the start codon.

STUDENT: There's a stop codon.

ERIC LANDER: There's a stop codon, at some point. And what's in between the start codon and the stop codon?

STUDENT: [INAUDIBLE].

ERIC LANDER: Well, no stop codons. A gene should look like ATG and a whole lot of codons without any stops in the reading frame, until you get to a stop. That's called an open reading frame, a long stretch without stop codons. So I could look for an open

reading frame.

So by an open reading frame, I mean a long stretch that starts with an ATG and then goes on and on and on and on with-- How frequent are stops? There are three stops out of 64. One codon in 20 is a stop, on average. So if I go 20 codons, I might see a stop, on average. But suppose I run for 100 codons, and there's no stop--- without a stop codon. That's pretty impressive, isn't it, if I can read 100 codons in a row, and I never see a stop, that's pretty unusual. So I say, that's an open reading frame.

That's one way to recognize the gene in there. The problem is introns. What happens if there's an intron? Yikes. Then it'll be spliced there. That'll be spliced out, but I won't initially know that, reading the sequence. And there could be stop codons in the intron because it's not part of the final message.

So I'm in trouble. So happily, in yeast, which has very small introns, not very many introns, I can actually almost get away by looking for open reading frames. In human DNA, this is kind of lousy. Well, it's really problematic because there'll be too many introns.

There, other tricks that get used-- I could make cDNA and compare it to cDNAs, which have already spliced everything out, and look for the open reading frame. Other tricks that get used-- I can compare it to the database of everything everybody has ever sequenced before and start looking for similarities. And today there are massive databases.

So many years ago, a postdoctoral fellow in my lab cloned a gene related to a human disease. And she didn't know what the gene did. And she found it. And it had exons, but she didn't know where they were yet. But she just took the whole sequence and said, this sequence here, is it similar to anything that's ever been seen before?

This was a gene that was in people who had a really severe form of dwarfism with twisted bones and things, called diastrophic dysplasia. And she put it against the

computer database. And the computer came back and said, the sequence you just gave me has a whole lot of patches that looks just like sulfate transporters in a fungus.

She instantly knew what her gene did. Because it turns out that bones have a lot of sulfated proteoglycans, et cetera, et cetera, whatever those are. And she instantly knew, because my sequence was similar to something-- it's a human sequence similar to something in a fungus that does sulfate transport, I've probably got a sulfate transporter. That's probably the basis of my disease. She took her cells from her patients, added sulfate, found that the cells couldn't take up sulfate very well, and bingo-- had found the cause of her disease.

One of the most powerful ways-- it's sort of Google, of course, right?-- it's Google before Google. You take your sequence and you Google it against all other sequences and see what it's like. And by googling all of life's sequences against each other, if somebody else has already solved your problem for you, you can find out about your problem. And it's just this wonderful network effect that is so characteristic of information technologies.

So anyway, you can do that by searching databases. And we will not, in this class. So you can write code for looking for open reading frames, you can search databases for similarities across organisms or within organisms, or things like that.

We won't here, but there are at MIT some great courses on computational biology that, for example, you can write algorithms for detecting these sorts of things. It's an interesting question. How do you write an algorithm for comparing two strings which might have insertions and deletions and changes and things like that? There's a whole rich field of computational mathematics associated with genome comparisons.

All right. So we've got it. Bingo! Now, here's our next problem.

Our next problem, we cloned the gene for beta globin from you. You were kind enough. You signed an informed consent allowing us to take some DNA, prepare a

library. We made our antibody, we found your beta globin gene, et cetera.

Now we're going to conduct a study of beta globin in a larger population. Maybe we're going to ask multiple people in the class, would they be willing to sign an informed consent to have their beta globin gene sequenced? It's an interesting gene. There are variants in it that confer risk of sickle cell. There are variants in it that confer risks of other things.

There are fascinating things about that gene. Maybe we'd like to see the beta globin sequence of many people. How do we get the beta globin sequence of a second person?

Well, how'd we get the beta globin sequence from the first person? Took DNA, cut it up, cloned it in our vector, spread it out on the plate, washed over the antibody-- actually, we took cDNA-- washed it over, et cetera, et cetera, et cetera. It's a lot of work.

If we wanted to do 100 people in this class, do we have to get DNA from each of you, prepare a library, maybe a cDNA library even, and do the same exact process to discover your beta globe gene? Or is there any way where after we've done your beta globin gene, we could now do everybody's beta globin gene a lot easier?

STUDENT: [INAUDIBLE]

ERIC LANDER: Actually, I do know the sequence. That's the thing that's different is having found it once, I know the whole sequence. The question is, can I use the sequence to save me the trouble of making an entire library of everything? How can I use the knowledge I've just gained to make it so much easier?

Well, the answer occurred to a chemist working at Cetus corporation in the mid 1980s. He was driving along, and he was thinking about this very problem and thinking about sequencing and how they do it. And he had the following thought.

His following thought was, suppose we've got the whole human genome. There's a whole human genome. And I'm just going to melt it for you, for a second, into two

strands. And suppose we've already discovered the beta globin gene. The beta globin gene is right over here, it turns out. That's beta globin.

Well, we know this whole sequence. This is total DNA. I haven't done anything right now. This is the whole human genome that runs over 3 billion bases. But it's the whole genome. I know the sequence, right? I could make a primer to that part of the sequence.

And suppose I just make a primer to that part of the sequence, throw it into your total DNA, and I add polymerase and nucleotides. Maybe it'll start copying. At some point, it'll fall off. But notice, I've made an extra copy of beta globin-- of course, mixed into the whole, total human genome. But there's a little bit extra beta globin now.

Suppose I also made a primer over here. I'd get that. I'd now have two double strands of beta globin, whereas before, I only had one. Let's call this step 1.

What do you think step 1 should be followed by?

STUDENT: Step 2.

ERIC LANDER: Step 2. Very good. Excellent. You guys have learned induction.

So let me melt the DNA and now throw back my primer. Actually, if you'll allow me, I'm going to make the two primers different colors. Let's make that one a different color. There we go.

So now what will happen is this primer goes here. This green primer goes here. This guy sits down here. And this guy goes like that. Well, that didn't come out very good.

I'll just draw it a little more clearly here. What happens is this guy will start copying this way. This guy starts copying this way. This guy starts copying this way. This guy starts copying this way.

Now, after step 2, how many copies of beta globin do I have? Four copies. What's

the next step?

STUDENT: Step 3.

ERIC LANDER: Step 3. Very good. No putting anything over on you. And after we melt the DNA and we add back the primers, how many copies of beta globin will we now have?

STUDENT: Eight.

ERIC LANDER: Eight, because it's doubling every time. Step 4? Step 10? 2 the 10th. 2 to the 10th, because we're doubling. 2 to the 10th.

Step 20? 2 to the 20th, which is about a million. Step 30? 2 to the 30th is about a billion. Oh.

After 30 steps, I have 2 to the 30th copies, which is about a billion copies. And at that point, the majority of the DNA in my tube is beta globin. The rest of the human genome is still there, but beta globin started out being one one-hundred-millionth of the genome. And I've just amplified it a billionfold. So it's now 90% of what's in the tube.

Pretty cool. This is like a chain reaction. You do it once, you do it again, you do it again, you do it again. You just throw in polymerase, and you run a chain reaction. This therefore is called the polymerase chain reaction, or as it is universally known, PCR.

That's PCR. That's the polymerase chain reaction. Kary Mullis, who invented this thing, won a Nobel Prize in chemistry for it. Because notice what he's just done. He's cloned your beta globin gene without cloning.

It's cloning without cloning. I didn't need any vectors, I didn't need any bacteria, I didn't need no nothing. All I needed was the sequence that I got once by cloning, and then I'm off to the races. I throw in two primers-- choop-choop-choop-choop-choop-- bingo!

Where do my primers come from? They're not in the catalogs. We don't keep all

that inventory. You just type them in, and they come to you the next day by an automatic synthesis machine.

So anyplace in the human genome or the yeast genome or any other thing that you want to PCR, just give me the two primers and piece it out. How do we do this? There are a couple of the details that I have to worry about here.

Cooking details here for the recipe. What I have to do is I have to take my test tube, and I have to heat it up so high that the double helix melts and comes apart, so that the primers can get in there. So I have to heat to 97 degrees.

Then it comes apart. I cool it down, I add my polymerase. I cool it, I add polymerase and nucleotides. And then it does its extension.

Then I heat it up to 97. Now the problem is when I heat it up to 97, you know what happens to my polymerase?

STUDENT: Denatured.

ERIC LANDER: It gets denatured, and it doesn't come. It's ruined. So what I have to do is I have to pop open my two-- Sorry, you had a question?

STUDENT: [INAUDIBLE]

ERIC LANDER: Why do I heat it up? There are ways you might be able to avoid it. But the traditional technique is you heat it up. But you're right, there might be other solution.

But now I'm going heat it up. And my polymerase gets denatured. So I have to pop open the test tube, throw in some more polymerase, let it do its work, heat it up again, pop open the test tube, put some more polymerase in. And it gets really boring. Every one of these 30 steps, I have to keep adding polymerase.

So the engineers in you will say, why don't we just design a polymerase that doesn't denature at 97 degrees? So we should go to an expert and say, please make us a polymerase that doesn't denature at 97 degrees, even that doesn't mind being boiled.

So what expert do we go to?

STUDENT: Bacteria.

ERIC LANDER: Bacteria. What bacteria do you think has a DNA polymerase that doesn't mind being boiled?

STUDENT: [INAUDIBLE].

ERIC LANDER: Bacteria that live in, say, geysers, hot springs. So you go to a hot spring. You go to geyser, you go to Yosemite, and you fish out some water, and you see what's growing there, and you find the bacteria growing there that has the name Thermus aquaticus. And you purify DNA from Thermus aquaticus.

Thermus aquaticus just goes by name TAQ, T-A-Q. You purify TAQ polymerase. And now, no problem. You just use TAQ polymerase, throw it in your test tube. And you go heat, cool, heat, cool, heat,cool, heat, cool, heat, cool, and you're all done.

You just put it on a little heating block. And the heating block automatically goes hot, cold, hot, cold, hot, cold, hot, cold. And that's called the thermocycler. The thermocycler does it.

And of course, nowadays, do you have to yourself personally go to the hot spring and risk your life fishing out the bacteria? No. Because it's in--?

STUDENT: The catalog.

ERIC LANDER: The catalog. Exactly. Very good. TAQ polymerase is in the catalog.

All right. I'll tell you a story. The statute of limitations has already expired, so it's OK. TAQ polymerase used to be very expensive. So we needed a lot of it in our lab. And we couldn't afford all of it.

So what we did was we just looked up the sequence of the TAQ polymerase in Thermus aquaticus, got primers, used PCR to get the gene for TAQ polymerase, and then expressed it to make a lot of TAQ polymerase. So it's kind of cool.

Anyway, that was about 15 years ago.

We produced in a few days what was then worth about $4 million worth of TAQ polymerase. [LAUGHTER] That was why we went to the trouble of doing it. We didn't end up getting in any real trouble about it.

OK. So now, why is this stuff cool? This stuff is cool because you're able to amplify tiny amounts of DNA. So if I want to purify any human gene now, and I know its sequence initially, PCR it out. No problem.

Suppose a patient presents with a bacterial infection. And you're a physician. And you suspect that there might be a specific bacterial infection or maybe a specific viral infection.

So applications of PCR. Application of PCR? Well, resequencing a known gene, yes. Resequencing beta globin.

But infectious disease. I have a patient. I think there might be a bacteria, there might be a virus in the blood. What do I do? Make primers, do PCR.

I have a detection technique. I can detect the presence of a viral infection or a bacterial infection. For example, HIV testing can be done by PCR. Because it doesn't take very much there in order to detect it.

Water contamination. You can test for bugs that shouldn't be in the water, by PCR because you don't need much.

How little do you need? Suppose I take a tube of DNA, and I start diluting it and diluting it and diluting it. How far down do you think I can go and still PCR back up, say, beta globin?

Suppose I dilute it so that there's only like 1,000 copies of beta globin left, on average. Can I still PCR it? 100 copies? 10 copies?

Suppose I dilute it so on average there's only one copy of the beta globin molecule there. Can I PCR it? How can I prove that?

An easy way to prove that-- I could do it statistically by just diluting it so that on average there's only one beta globin. But how can I get one copy of beta globin packaged up very nicely?

STUDENT: Order it.

ERIC LANDER: Sorry?

STUDENT: Order it.

ERIC LANDER: No. Can't order that. Where do you know that there's just exactly one copy of the beta globin gene? One human sperm.

Suppose with a micromanipulator, I purify a single human sperm, one sperm. It's haploid. It's got exactly one beta globin. Throw it in a test tube, crack it open, do PCR, it works.

That's how I demonstrate that a single copy is enough. I can make that work. Pretty impressive.

Not only that, that I can do it from a single copy in a sperm. If someone is doing in vitro fertilization-- remember, in vitro fertilization won the Nobel Prize this semester-- in vitro. Suppose a couple has a 1 in 4 chance of having a baby with some terrible lethal disorder. The couple might use in vitro fertilization to make multiple independent embryos.

The doctor, then, is deciding which embryo should we implant back in mom? Well, how could they tell at this eight-cell stage which embryo carries the genetic disease? Suppose this genetic disease they already knew the molecular mutation causing it. Pull off a cell. Pull off one cell, pull off one cell, pull off one cell. This is at the eight-cell stage, let's say.

If I pull off one of those cells at the eight-cell stage, does that mean the baby doesn't have an ear or an arm or something? No, it doesn't. Because at that stage, none of the cells have taken up any identity. It regulates. There's no problem. It

turns out you can pull off an individual cell, and it has no impact on the embryo.

And I do PCR. And I can figure out that that one carries the severe genetic disease that's going to cause the baby to die at five months. And the couple says, we're not going to implant that one. We'll implant the other ones. That's called preimplantation diagnostics.

Or suppose somebody's being treated for cancer. And the cancer cells that had previously been there are no longer detectable. The drug therapy has apparently killed this blood cancer that somebody has.

Maybe they have a cancer of the blood. Now what I'm going to do is monitor that patient every several months by getting a blood sample and seeing if the distinct mutations that were present in the cancer were there. And I can begin to see if that's coming back, if those cells are now recurring.

It's an incredibly sensitive technique. And then of course, where does this stuff get used that you guys all surely we know about? Forensics. *CSI* and all that kind of stuff.

If I lick an envelope-- I used to say, licking a stamp, but stamps just peel off these days. But people still do lick an envelope sometimes. If you lick an envelope, more than enough DNA comes off when you lick the envelope that you can use PCR to determine who do the licking.

ERIC LANDER: You can. It works. All right. So that's PCR.