

MIT OpenCourseWare  
<http://ocw.mit.edu>

7.014 Introductory Biology, Spring 2005

Please use the following citation format:

Penny Chisholm, Graham Walker, Julia Khodor, and Michelle Mischke, *7.014 Introductory Biology, Spring 2005*. (Massachusetts Institute of Technology: MIT OpenCourseWare). <http://ocw.mit.edu> (accessed MM DD, YYYY). License: Creative Commons Attribution-Noncommercial-Share Alike.

Note: Please use the actual date you accessed this material in your citation.

For more information about citing these materials or our Terms of Use, visit:  
<http://ocw.mit.edu/terms>

MIT OpenCourseWare  
<http://ocw.mit.edu>

7.014 Introductory Biology, Spring 2005  
Transcript – Lecture 27

There was a little confusion with dideoxies in one sense, and some of these things like the PCR you're going to have to sort of sit down and actually think about it, but the principle of the dideoxies, if we were making a chain of beads that had a hook on one end and a little hole on the other, and we were joining these things together, we could make obviously, a chain that went on. And then we could hook another one in, and so on. And if we had a bunch of beads like this every now and then we threw in a very small number that didn't have the hook on the end, any time this particular chain were elongating and we put on one of these things, the chain would stop because you haven't got any where to join onto it.

You added dideoxy into a polymerase reaction. A chain that gets the dideoxy doesn't have anything to join in the end, and that will stop. If we only added this, the entire reaction would stop, and everyone would come at the first time a dideoxy got incorporated.

The trick is to put a little bit in. So, a few of the molecules stop. Everything keeps going. The next time a dideoxy gets incorporated, the chain will stop there. And out of this, you will generate a family of polymers that are of different lengths. Each one will terminate with a dideoxy nucleotide, and if the dideoxy nucleotide we used in that particular reaction was, let's say, dideoxy ATP, that means that an A was the last nucleotide added to every one of those. And we can separate these on the basis of size. And if I ran them out on a gel, I'd see something like that.

And that would tell me that when that polymerase was coming along, that was the first time it saw an A. A few stopped there, polymerized a few more. Then it put in another A, put in some other things, and so on. And that by itself wouldn't tell us the sequence. But if I did that reaction four times in a row, then I could tell. In the old days, they didn't used to use dyes.

We just did P32 on it as a label, and then you'd run the four reactions side-by-side. And this would be with dideoxy ATP. You'd see a pattern like that, and maybe with dideoxy TTP, you'd see something like this. And when you got the rest of them, you'd kind of end up working out what the sequence was by looking across the four lanes. This business of using the dye is just one more step up in the engineering side that enables the thing to be done automatically. And it's pretty well explained in your textbooks.

OK, PCR, someone was confused as to why we didn't just let the cell do it. Well, the cell does a great job, but if you are a molecular biologist trying to understand the basis of life or if you're a biological engineer, and you want to produce something, you need to get hold of a particular piece of DNA. Or, if you're a forensic investigator, and you've got a tiny, tiny sample of human DNA, and you want to know whose it is, you have to make more of it. And, that's what PCR is all about.

So I'm going to switch just over to the net for a minute. I think this first site, I just want to show you something, how somebody functions in a lab now with all these

genes out there. And then, I'm going to show you a little animation for PCR that will help. So, if you just go to Google and type NCBI, that's the National Center for Biotechnology Information. And, the Dolan Learning Center is a center that Cold Spring Harbor Laboratory has set up to teach people about DNA.

So let me just see here. So, let's just go to, OK, let's use, whoops, this is going to seize on us. OK, let's find how it happens. OK, so here's this National Center for Biotechnology Information. There's all sorts of things you can search for, and I'm not expecting you to know the site. I just want to sort of give you a demo. If I was sitting in my office, this is the sort of thing I can do easily. Rather than sequence looking for DNA sequence, I'm going to look for the translated sequence of the protein that's encoded by the gene where the computer's gone through and used the genetic codes tell me the sequence of a protein. I told you about sequencing a mismatch repair gene back in the 80s. It was called MutS, and I'll put in Walker GC, and probably hopefully get us to the thing. And there, the very first hit is DNA repair.

Protein MutS salmonella typhimurium, that's the one I sequenced. So I'll just go to that. It has various ways of displaying the sequence. I'm going to switch to Fasta, which is a very easy way to see it. Now what you see is the sequence of the protein using a one letter code, or one letter stands for each amino acid. K is lysine. A is alanine, and so on. I'm just going to copy that, that piece of sequence. OK so that's the bacterial gene for mismatch repair.

At the time I put that in the database, there wasn't anything else like it, except for the gene that was Streptococcus pneumonia. But I found out, someone else is sequencing by phoning around in the field. I'm going to go back to the main site and I'm going to use a program called Blast, which lets you search the entire database. I'll use a protein blast. I'm going to take a protein sequence, and I'm going to ask what else is out there in terms of protein sequences? I'll paste in this bacterial sequence, and then I'm going to, if I can, manage this thing.

Let's see if I can get myself down here. OK, over here I'll probably do, OK, so I'm going to limit it, let's just search the human genome. That's all we've got to do. And what did I have to do to get this thing to fit? Which button? Go to the right. Can you just come up here for a second to help me get this set up? I'm computer limited here apparently. OK, that one, OK, great, so why do we not try this again? PAUSE] Sorry about this.

We'll see if we can get this thing to go. I have what? Yeah, that's OK though. That should be fine. Try again. Let's see if I can get the thing to work. OK, so it's got it now. It'll tell me. It's searching all the sequence that's out there. There's just an unbelievable amount of sequence. That's just how long it took. It's showing me here a diagrammatic representation of the things. Then I can see that the very first it was MutS homolog three for humans. And if I go down here, we can actually see on alignment of the bacterial gene on the top line, and the lines below is the sequence of that particular human homolog. And you can see in between all the things that are in common, and particularly down at the C-terminus of the protein, you can see there's very strong conservation.

You may not think that that's impressive, but remember for every one of those positions, there's 20 possibilities. So, if you get that many in a row, that's the same gene basically. And when you take the structure, the structure's going to be very, very similar. And it does mismatch repair in both.

Just to try and give you an idea of how you do sequence now, because with all these genomes done, you do the vast majority of it by computer, rather than some other way. I want to take you and show you this. DNA learning: if you go there, the second thing is a set of animations. Go to the animations. There's one on polymerase chain reaction. And, I'm going to just show you this because this is a nice little, let's see if we can get this thing to center. OK we'll have to see whether this is going to work. OK, so this is the principle of, you can go do this at your leisure, but the idea is to heat the DNA up, the strands come apart. Then, were going to take these two little primers, not promoters, which I think someone was confused about, a little piece of DNA that complimentary, and anneal them.

Then we added DNA polymerase. You know what happens then. We extend those primers. That was the first cycle. Do the same thing again. Go to the second cycle. This is what I was drawing on the board the other day. Now are going to denature the DNA. The strands come apart. Let's let the polymerase extend them. Let's go to the third cycle, denature the DNA, anneal the primers, extend the primers, and now for the first time, we've got what we were shooting for. We have a double-stranded copy of just the DNA that was defined between those primers.

OK, I think this actually, I'm going to go back one. Oops, OK. If you go, then, to the amplification graph, what they're doing here is they're showing you what happens as you do successive cycles. So, at the first, oh, it's down here. Just a minute. If we do the first cycle, we end up with two DNA copies. That's just plotting what I showed you. The next one: we have four. We haven't yet got to this target sequence. By the next cycle, we now have two copies of the target sequence plus these other things.

But if you keep going, let's say by the time we're at seven cycles, the number of targets is up to 114. The number of DNA copies is 128. But, if we keep going like this, we'll find out that the target copies become the vast majority of the sequences that are in there. So, by the time that you're up in the 30 cycles, or something like that, there's only a handful of the original things, or almost all, and that, I hope will help some of you who might have had problems with understanding the PCR. So, what I'm going to do is tell you a few more things about what you can do with recombinant DNA, this recombinant DNA technology, because it's just so powerful.

And I can only sort of give you a few ideas, and show you a few variations. But, most of these things are just taking principles that you've already learned as part of the basic biology I'm trying to tell you, and then using them like an engineer to achieve some applied purpose. For example suppose I wanted to produce a human protein, and try and produce it in a bacterium.

That would be great. I could take the one gene. I could grow a fermenter load of E. coli, and if I got it right, then I'd be able to make a lot of this protein instead of trying to isolate it from some human source or something like that. There's a couple of problems. We talked about them. One is the problem of promoters. Another one is that human DNA would have introns in it.

And, bacteria doesn't recognize the human promoters. It wouldn't start to make an RNA in the right place, and it doesn't know what to do about splicing out the intron. So, let's address the intron first. There is a way of handling that that's quite easy. And that's what's called to make cDNA library. So, if we have DNA, and then we get the RNA, we get the RNA copy [SOUND OFF/THEN ON] including these intron

sequences. And then what happens, this is RNA splicing that we talked about. And what we get out of that would be an mRNA, in which the introns have been removed. So, eukaryotic cells, my cells, know how to express the gene, so they make the RNA, they know how to get rid of the introns.

So, if I were to isolate a messenger RNA. That's been spliced from me, or you, or anything, what we would have is a population of RNA molecules that don't have introns anymore. Anybody remember any way we could get from RNA back to DNA? Reverse transcriptase. So, if we used reverse transcriptase, that protein that David Baltimore discovered and viruses, and which retroviruses use, now we would have a single stranded DNA copy of the mRNA. And then, we could use an ordinary DNA polymerase to get ourselves to double-stranded DNA.

We'd be doing, in essence, exactly what one of these retroviruses does. This would give us what's known as a cDNA library, where the genes don't have introns anymore. So, if I wanted to get at one of my proteins, one of my genes, and think about expressing it in *E. coli*, what I would do is go looking in a cDNA library using a sort of approach as we've done, trying to find my gene of interest, because if I use the cDNA library now it would just be like a bacterial gene. You could see the ATG start.

You could get out your handy little genetic code, and you could walk along, and read out the sequence of the protein. So, that's part of what you need to do if you wanted to make, say, a protein inside of a bacterium. The other one which we talked about was since the promoters are not a universal language, what *E. coli* RNA polymerase sees is different than what human RNA polymerase sees as a start site. I would have to add in a promoter that would drive the expression of this open reading frame if I wanted it to work in *E.*

*coli*. And that's fairly easily done, too. A general thing that's for this is called expression cloning. And it would be more or less the same idea. We'd have a vector that had a cloning site. It has an origin of replication, and maybe there's a selectable marker such as the drug resistance. That's the basic kind of factor that I talked about before. However, if I clone in a piece of DNA into that, it has to have a promoter that can be read in the organism working with, because it's just out whatever nature gave it, whatever promoter that would be in front of that. But, if I were to now into this vector put an *E.*

*coli* promoter right there, now, if I just downstream of that put any open reading frame, human protein lets say, which we've gotten rid of the introns, the human genes minus its introns, which you got from the cDNA library, now when the bacterial polymerase came along, it would be copying, making it a messenger RNA for human protein, and we could get it out of that. And the beauty of that, suppose we took the front part of the Lac operator, we would have a regulated promoter.

It would be just everything we studied about Lac if we were to starve it for, you know, we have to get rid of glucose, and then if we added lactose or some kind of synthetic inducer, we can turn the promoter on and off. So, you could grow an entire fermenter load of bacteria without expressing the gene. And then, once you had the bacteria all grown up, you could throw in something that would normally induce the expression of the Lac regulatory system. And now, instead of making beta galactosidase, instead it would make the protein that you are interested in, you with me? It's very pretty.

And in fact, so much of what you can see in this is, these really basic studies, since the Lac system was one of the first to really be worked out in detail, we use its parts. And, there are many vectors around now that have exactly that. They have the Lac promoter, and you can turn things on and off in a regulated way, so not only provide a promoter that works in the organism, but it also gives you a measure of control.

There's another very cute trick, and what we've done sort of here is we took, say, the promoter for Lac in the regulatory region. I'll use R to stand for regulatory region, and then this would be the LacZ sequence. What we've really done, is we've taken a gene from somewhere else, let's call it gene X that had a promoter from gene X. And, in essence, is cutting each of them here. And, now we take the regulatory promoter region from Lac and we put down below it gene X. And, now we've got this gene whose products we're interested in producing in a fermenter under the control of the Lac operon. Well, there's another kind of thing we can do.

We can do the other way around. We could take LacZ, which makes beta galactosidase. We could put it under the promoter regulatory region of gene X. Well, what will happen then? If that construct is sitting in a cell, anytime that the cell decides to make gene X, instead it will make beta galactosidase, which is really easy to assay for. And, this sort of strategy, you'd use something like LacZ as a reporter. In this case beta galactosidase synthesis, which you can assay for, reports when is the promoter of gene X is functioning.

This reporter gene now has the regulatory characteristics that are imposed upon it by that particular promoter. So, there's this picture that I've showed you, this little movie I showed you early on. You've seen it a couple of times. In this case, the reporter is green fluorescent protein. What Barbara Meyer, who made this particular construct, did was they took the gene for green fluorescent protein which started out in a jellyfish as you may remember. And, the protein folds up, and ends up being fluorescent.

So, we can tell when it's expressed very easily. And in this case, you'll notice not all of the genes in the whole worm isn't glowing. And so, it's under the control of the promoter regulatory region that is expressed only in specific body parts. And so, you can see where that promoter is working by just looking at the worm. In the case of something like the mouse that we talked about, it's a pretty uniform expression at least in the skin. So, that was probably, in that case, the green fluorescent protein was probably put in a promoter that's expressed in probably most of the cells in the body, at least certainly all the ones in the mouse cell. I don't know the details of that. Ditto over here. It was probably something that was expressed in most of the body cells, but you also could have put something that was just expressed in some very little bits.

So, depending on how you do the construct, there are a lot of different things that people can do in this sort of thing. OK, one more category that comes out of the sort of thing, is if we have a gene of some type, I don't know what it does but I'd like to find out. You know, at least budding geneticists, know what we'd like to do.

We'd probably just like to disable that gene very specifically, and then look at the live organism to see what happens. And, the principle is the same whether you're doing it in E. coli or a mouse. It gets a little more complicated for technical reasons doing it in a mouse. But the idea is exactly the same. And here's the strategy. So,

we'll just take a piece of DNA from the organism. And sitting at here is this open reading frame that we've seen. We don't know what its function is. We think if I could knock it out, get rid of its function, I'll look at the organism. Maybe I can make a guess then. So, if we were to cut the gene somewhere with a restriction site, and then we were to take, for example, a gene encoding a drug resistance or something like that, and insert it at that point, what we would end up with is this piece of the organism's DNA. The first part of gene X, then a drug resistance, then the last part of gene X, and some more sequence from the organism.

Now, this would be, we'd have this in a test tube. We could do it by the kind of recombinant DNA manipulations that we have. And what would happen if I were to put, now, let's keep it with bacteria where it's easy to see. If I were to take that piece of DNA, put it inside a living cell, what's going to happen? Well, let's make this, say, here's the end of it.

That's all we've got. Well, inside the living cell, we of course have the entire genome. And then we come to this part. We have gene X. Then we'd have this going. That would be the whole thing. Well, this particular piece doesn't have an origin of replication. It's not joined to a vector. It's just sitting there. So, if the cell divides, it's not going to get replicated. So, if I select for a drug resistance that's on that piece of DNA, unless something happens I'm not going to get a drug-resistant bacteria. But you do know a way that this thing could join to an origin of replication. It could join to the origin of replication that's on the bacteria chromosome.

And, the way to do it would be by undergoing genetic recombination over here, because this DNA is exactly the same as on that side, and over here it's the same thing. This DNA is the same as that side. So, if that genetic exchange happened, what would happen, even if it happened rarely, was this piece of DNA would replace the piece of DNA that's in there. I'd be able to tell it was there because I'd just select for drug resistance.

And even if it only happened only one in 500,000 cells, it wouldn't matter because up would grow the colony that now has the drug in the middle of gene X. Gene X is gone, and I could look at the organism if it's alive and see if it has a phenotype. If it's an essential gene, that strategy obviously won't work, and when people do the more complicated thing of doing this kind of experiment to make a transgenic mouse, it takes about a year to go from our DNA manipulation all the way to the live mouse with a disrupted gene. And sometimes what they find after spending half of your PhD. is that that was an essential gene. And, there's no live mouse, or it made it two days into being an embryo and it tanked at that point. But, this again, you could see, we talked about going back and forth between gene, protein, and trying to figure out function.

All I can sort of do is give you the flavor of what's going on. But one sort of overarching thing I hope you remember going through this is DNA sequencing, PCR, all these kinds of manipulations we're talking about are just exploiting these basic cellular components that we learned about studying, how does DNA replicate? How is information coded? How do genes get expressed? How does genetic information gets sorted between cells? It's simply applying those relatively well understood tools, or sort of biological principles and parts that we learned about, and now using them as tools in an engineering way, and have just completely transformed the way biology has been done in the last couple decades.

And it's just, as I say, things are changing so, so fast. It's almost breathtaking. So, the last little bit of sort of technique oriented stuff, I just want to at least make sure I've mentioned what are called microarrays. You often hear these referred to as DNA chips. The principle here is this is a way that lets you ask, not only whether one gene is being expressed or not under a particular condition, whether its RNA is being made, and in most cases that means making protein, or whether it's off, or whether it's at some intermediate level.

A microarray lets you do that experiment with many, many, many genes at once. And here's the principle. You take some surface, and there will be a bunch of, if you will, sites on the surface, on this chip or whatever. And, what will go to be attached here would be a little piece of DNA from gene one. I mean, let's say, maybe a hundred nucleotides: that would be far more than enough to make it absolutely specific that they could only hybridize to a messenger RNA from gene one, and not from anything else.

And, this one, then, would have from gene two, this one, from gene three, and so on. Then, if we were to take a messenger RNA preparation from, say, an organism if it's a little one, or maybe a tissue, or something like that, anywhere you could isolate RNA from. And then, we'll label it in some kind of way, and we can label it radioactivity, we can label it with dyes. It's usually done with dyes, and there are a variety of variations on this. Those are sort of technical details how to do it. But here's the principle. Let's just, for the moment, just consider that it's got a label on it.

So if we take the messenger RNA, and take this little chip that has samples in the extreme, it could be a sample of every single gene that's in the genome of that organism, we take this labeled RNA, actually what we would usually do is to use this to make a labeled cDNA preparation, which would be a copy of each one of these things. That the technically easy way to get label into it.

But what we do have is if the gene was on, its messenger RNA would be on, and we'd have a bunch of stuff corresponding to gene one that had label on it. And if we give it a chance, that will, then, hybridize here. And there would be some way of detecting this label. If gene two was off in that sample, there won't be any hybridization. There won't be any signal. You can sort of see in principle what you're doing is you're interrogating each gene in the extreme, each gene in the organism under some condition. Is it on? Is it off? If you did various samples, and you could see maybe it's in an intermediate level, and so on. So, the chips look like that, 50-100,000 genes perhaps, something like that.

These things are really small. Here's sort of a display of a simple one, and this is one where they're taking RNA. The samples are from two conditions. One's labeled with a dye that's green, and one's labeled with a dye that's red. And, if you get equal amounts, it looks yellow. So, they mix the two things together, and if the gene is the same under two conditions it would be yellow. Under one condition, if the gene was on in condition one than it would be green, and off in two, and back and forth.

So, without trying to get lost in the technical details right now, which doesn't matter, the principle of this thing is that you can take, you can sort of, by making a preparation of RNA, then you can use these DNA chips and say, is each gene on and off? Or if I switch conditions, who comes on and off? So, it's a little like, I think of it this way.



It's like having, all right, who's on today? And a number of hands go up, or something. And the rest of you would be off. But, come back on Monday, and I say to the something or other, and a different set of you would put up hands. And what I'm kind of looking at are the changes between that. And the sort of thing where this has been so powerful, for example, is there are kind of cancers for which there is a treatment, but it was only 20% successful. And, when people started to study these cancers and then looked to see what genes were on, what they realized was even though physicians had given these cancers a particular name, if you looked at which genes were being expressed, they fell into two classes, class A and class B.

And what they then realized was that the treatment they were using was 100% effective of tumors of class A, and wasn't doing anything for the tumors of class B. The physician couldn't tell the difference between these two types of tumors, but a microarray can say it. And once you have that kind of insight, if somebody has a tumor of the type A, you can use the treatment. It's very effective. And of course, when people are trying to do now would be to develop an equivalently effective treatment for something for the tumor of type B. So, again we have so little time in this class.

I could go on basically for ages. There's the output of the real sort of DNA chip. You can see things are very dense, and the great cleverness in doing these things, people now use the technology that goes with LaserJet printers to actually synthesize little pieces of either DNA or proteins starting on a little spot on each membrane, or on the chip, or whatever. You put it on one nucleotide and then you put it on the next, and the next, and the next.

You could sequence it using technology that's already around for inkjet printers at that kind of thing. So here you've seen a fusion of different types of engineering. OK, the last thing I'm going to tell you about is a little bit about the immune system. We've run into this. This is a movie that some of you liked, got the biggest aw I think of the last part of the course anyway. But what we are seeing here is a white blood cell pushing aside from red blood cells, which are stationary, and chasing a bacterium. It's obvious that it can trace it. It's able to recognize some things, and at some point, then, it took it up. The principle of what happened there was this white blood cell had a capacity to recognize the bacteria, then bind to it. And then, its membrane, this is the membrane, and this is a white blood cell.

There's many types of these, and it pinches the membrane off. So you have the bacterium. This is the bacterium. And, it's inside a little membrane compartment, as if the bacteria is in a little soap bubble. And the principle of what happens is the white blood cell has another soap bubble that's full of poison. And, if you took two soap bubbles and push them together, you know what happens. They'll fuse, and you'll get a bigger soap bubble. And that's, in essence, how these white blood cells normally kill bacteria. They would bring together these two compartments. Now you have a bacteria and a poison within a white blood cell at the bacterium, would get killed.

And we talked about how bacteria fought back. That was a streptococcus that has a capsule. And the capsule, by having polysaccharide on the outside, prevent the white blood cell from being able to grab hold of some feature of the bacteria, that starts this process of killing it. And when I told you the story of how DNA was found, it was people studying pneumonia. If you remember, it was streptococcus. If the

streptococcus had a capsule, and the people would get very sick. And after five or six days, there would be a crisis where they either lived, or they died. And what would happen in that time is that, the last thing I'll tell you about, what's called the adaptive immune system would have generated special recognition molecules called antibodies that would have learned to recognize the capsule of that bacterium. And once those were there, now those white blood cells would be able to capture the bacterium, because the antibodies give it a hand in recognizing there was something there that needed to be killed.

The way this adaptive immune system works, it's almost like science fiction. And I'll tell you the molecular basis of it. The key insight came from Susumu Tonegawa, another member of the MIT faculty in biology, and also runs Picower Center, who got a Nobel Prize for understanding the basis of the diversity of the immune system. What I just want to do for the moment is just sort of point out the key features of what's called the adaptive immune system.

And this is one of the reasons that we are able to live. And even though we get sick from time to time, and we've all had one thing or another get us for a little while during the semester, the reason we aren't sick all the time, and the reason we recover when we get sick, as we have what's called an adaptive immune system. What happens when people get infected with HIV virus is the cells that it lives in and destroys are key players in your adaptive immune system.

And to some people don't die from the HIV infection itself, they died because lots of things that we just, bacteria or fungi, whatever things we just have on us and we live with all the time suddenly become killers because you lack the immune system that fights them off. So the so-called adaptive immune system is absolutely amazing.

So several features, it's got an incredible diversity, there is a general word that's used to describe some sort of all sorts of chemical entities, and it's called an antigen. So anyway, this can recognize many, many what are called antigens. And at the moment I think you can just think of them as some kind of chemical entity. It could be a carbohydrate. It could be a little piece of an organic molecule. It could be a few amino acids on a protein. But it's something that's potentially capable of being recognized by your immune system. So, there are many, many, many things.

And the amazing thing is I can go into a lab and synthesize a molecule that's never been seen on this Earth before and challenge somebody with it, and you'll produce an immune response that will be mounted against that even though it's never been on Earth before. It's also the specificity. It is completely amazing. If I were to take a protein, and then, let's say, put on a phenyl ring with a methyl there, inject it into someone, the immune system would figure out how to recognize this thing with the phenyl ring, and the methyl. But, the response it generated, it would see this but not, let's say, that if I wanted to get an immune system, something with methyl here, I'd have to put that into the organism and let the immune system figure out a response.

So, the specificity is at the same kind of level that you are used to here. If restriction enzymes can read different sequences in DNA or protein can tell one optical isomer of a small molecule from another, it's this fitting of complementary shapes. So, with the immune system is all about is figuring out how to get a complementary shape somehow that's able to recognize essentially any kind of chemical shape and structure you can think of.

It's just mind blowing. And, you could already see right from the beginning, where the fundamental problem people could see from the beginning. It sounds like we would need a genome that's infinitely big, full of things that are ready to recognize anything. And so, one of the real surprises is, and now we know, is there is 20,000 genes are so in the human genome. There can't possibly be a zillion genes, each one specific for one of these structures. There had to be some underlying principle that we had to learn. And that was one of the big challenges in the immune system for a long time.

Another one was that if you have an organism that has this capacity, and you could recognize it, why don't you do yourself in? Because you yourself are full of entities that could, in principle, generate an immune system. So, one of the other things the immune system had to deal with was avoiding self recognition. If you're able to recognize anything, how do I not avoid killing my own selves? So, that is another really fundamental problem in this immune system. This is exciting. We can all stop and watch, but I think I'll just try and keep soldiering along for the last minute or so.

So, one other feature that was interesting about the immune system, is it has a memory. And I'll tell you more about antibodies at the beginning of next lecture. But, these are the kinds of molecule that's able to recognize these different entities. And you've all heard the term in your ordinary life. But, if we look at the level of antibodies that are made in the body, if the first exposure to an antigen that would get some kind of response that comes up upon the first exposure, and this is time here.

If we let there be some delay, it could be even into years, and then we get a second exposure, the antigen, the response is much higher. And this could be a log scale. So, it could be dramatically higher. So, what was the basis of that? How does it work? You see right there the principle of vaccination in the sense that if you ever got chickenpox as a kid, your body has learned how to make antibodies. So if you ever see it again, it mounts a really big immune response.

If you want to have a disease something like tetanus that you haven't seen, you go to the doctor and they squirt and a bit of the stuff that doesn't make you sick, but it gives you the initial immune response. Then, if you ever step on a rusty nail, you get a very powerful response against tetanus. And that sort of the underlying principle of vaccination is this concept of memory. We'll pick that up on Wednesday. So, have a great Patriots Day weekend.