

So in this last lecture, what I'd like to do is I'd like to now begin to talk about genetics. And when we talked early on in the semester I was showing you this way that you can study biological function. Biochemists, for the most part, study how proteins work. And what genetics does, steady function, it's enable one to discover genes. And the molecular biology that we're going to be talking about in the back part of the course has been really totally amazing because it's allowed one to go back and forth between genes and proteins, something that used to be very, very hard until recombinant DNA came into the picture. Now, up until now I've sort of mentioned genetics as we went along. But we haven't really talked about it as a discipline and the kind of power that this experimental approach has. Thank you. I could have done this without notes but it's easier to have them. I just have my little help here. So I know some of you have written in your comments I didn't think there was going to be so much chemistry. Well, we had a little bit of a bout with chemistry. And then some of you found that we were going away from bonds and thinking about 3-dimensionality, and that was another kind of thinking. And then sort of trying to work through the genetic code is almost another. And I think you'll find here, as we start to go through genetics that you have to stretch your brain in another direction. And part of one of the really interesting things about biology today is you cannot just sort of think about it in one kind of very comfortable way of thinking and really take advantage of everything that's out there. You need to be able to talk in multiple languages and multiple disciplines. And so today I want to just kind of briefly give you an introductory sense of how genetics, that I hope will give you, let you see the power of it by very, very simple techniques. And remember what you're doing here is you're studying variants of the living organism. So if you see something that's wrong with it you know it's important, but you know the rest of the job is to figure out what's wrong. And I've sort of given you and used some examples. When we talked about the streptococcus with the smooth colonies, I talked about how people had noticed rough colonies. Well, those sorts of things had been a change in the DNA and they weren't making these polysaccharide capsules anymore. Or when I talked to you about the mutant bacteria that had lost mismatch repair, they had a high mutation frequency compared to a wild type. So up until now I've sort of mentioned them, but we haven't really talked about how one goes about studying things in a systematic genetic way. So let me begin with just a few definitions. I realize this is a little dry but we need to make sure that we're all on the same page in terms of the language. So a mutant is a variant of a normal organism. It has a change in its DNA and the kinds of mutants can vary all over the map. For example, if we had an E. coli that was, we might call it penn-resistant, that could mean it was resistant to penicillin and would grow in the presence of the antibiotics. Whereas, ordinary ones would die. Or if we had what we call a his-minus mutant broken in the biosynthesis of histidine won't grow -- -- unless you add histidine to the medium. Those are a couple of simple examples. When we start right after lecturing, when I come back, I'm going to begin with Mendel, which is the more classic and usual way of teaching genetics. And he used various traits of peas. Yellow and green seeds and wrinkled and round. And you can see what he was working with there. Another favorite organism for genetic study has been the fruit fly or drosophila. You can even see that they have red eyes. If you take a close look next time if one lands on your sandwich or your drink or something in the summer. But you can see, for example, people were able to get white mutants that have white eyes. There's something broken in making the pigment that's going to make it red. Here's an even weirder one. This is a single mutation in drosophila. It affects a developmental process instead of something else. And what happens is this is a drosophila head. And normally there are antennae that grow out of the top of the head. And you can see what's happened in this mutant is there's a pair of legs growing out of the top of the head. It's just a single gene that's been changed, but it's a gene that plays a role in the developmental program, the patterning that makes certain cells become specialized to become certain other things. So it's just to give you an idea. And I showed you these. We've talked about the xeroderma pigmentosum, one change. A person with this has got a gene that's broken in dealing with damage from UV. Or I've given you an example of what they call Werner syndrome where a change in a single gene, it's actually a kind of gene whose proteins are involved in unwinding DNA, of all things, can give you this premature aging phenotype where the woman looks normal at 16 but older at the age of 48. So all of these are, if you will, mutants that are a variant of a wild type. And then this is something people often get confused between these two terms. The mutation is the actual change in the DNA. Now, one other thing that's sort of implicit in this definition is that we know what a normal organism is. Well, if you look around this room, what's a normal human? Well, I'm OK but I don't know about the rest of you. I'm sure we all feel that way, but we have a lot of variation in it. So, to some extent, it's an operational definition. And most often it's applied, if you're studying something like drosophila or E. coli, someone has been propagating a particular bacterial isolate in the lab for a long time. And you call that one the wild type, even though, in fact, in the case of an E. coli, it may have changed a bit so it likes to grow in the conditions that it finds in the lab. But this sort of normal is an operational definition. Now, another important term that geneticists use all the time is the phenotype. The word phenotype. And that's the ensemble -- -- of observable characteristics -- -- of an organism. For example, in these resistance to penicillin would be an example of a phenotype. A mutant that had acquired the resistance would have the phenotype of being resistant. Sometimes these phenotypes can be a little subtle. A conditional phenotype would be a phenotype that you could observe under one characteristic and not in another. For example, a temperature sensitive phenotype -- Whatever it might be. For example, it could be wild type at let's say 30 degrees but mutant at 37 degrees centigrade. That may sound like a sort of fiddly little thing. Why am I telling you about a conditional phenotype right off? Well, suppose you wanted to study a DNA polymerase. be it E. coli or me, and it's the one that replicates your DNA? If we get a mutant that just kills, knocks out the activity of the gene we have no living organism to study, we cannot do any genetics. But you can work around that. You can do genetics with essential functions if you get a mutant where you see the mutant phenotypes, let's say, at the high temperature but not at a low temperature. And what that kind of thing usually comes from is a change in the protein where the protein has one amino acid changed to something else and it folds up. and at the lower temperature it's able to fold up and do

its thing. But it's not quite as stable as the original protein. And if you raise the temperature it unfolds a bit, and once it unfolds it doesn't work or it gets degraded or something like that. OK. So that's phenotype. The genotype then refers to the state of the organism's genetic material with respect to whether its wild type differs from it. So let me just put that down. So this refers to the state -- -- of the organism's genetic material -- -- with respect to whether it's wild type or it's mutant. There's a key distinction, even though they sort of sound the same. This is what you can see if you look at it but this is what's actually changed. The genotype is what's actually changed in the DNA. And there's a caution. If it's something like a bacterium it's fine because it's only got one copy of every gene. If I knock out a gene for making histidine it cannot make histidine. The genotype and the phenotype are the same. But if you have a diploid organism, which is the kind of organism that we are and peas are and fruit flies are, there are two copies of most genes, one from mom and one from dad. And the only exceptions are the ones involved with the sex chromosomes. Then you can get to something else, a more complicated situation. Because let's say we have a usual thing where both copies of the gene are plus, then the phenotype is wild type, but if we were to break one of those copies by a mutation this one is still wild type. The phenotype is still its wild type. You cannot tell from looking at it that this one is different than that one. And one of the really brilliant insights that Mendel had when he was looking at peas was he would take things that were wrinkled and round and crossed them and he'd see mixtures of things. And he realized some of those things in there weren't the same as either the parents but they resembled one of the parents. And I'll take you through that in the next thing. So it's just a caution at the moment that you have to be careful that phenotype and genotype are not always the same. And then I've used the word gene. That's the discrete unit -- -- of genetic information. We talked the other day about the lacZ gene that encodes beta-galactosidase. OK. So I think what I'd like to do, if I tell you, for example, we made a lot of mutants of E. coli that were broken by the biosynthesis of histidine and gave them to you, if you were a good biochemist you might be able to work out how they differed by studying them biochemically. And let me just sort of give you a sense of how that would work so you can see histidine or any of these amino acids. They're sort of complicated and they have to be built up by a sequence of biochemical steps with one enzyme catalyzing each step in the pathway. And when people started out trying to study that kind of thing they didn't know how it was made. All they knew was what the end product was. And furthermore it was more of a biosynthetic challenge even then trying to work out the steps of glycolysis because actually quite a reasonable proportion of the protein in any cell is devoted towards making energy. So, relatively speaking, there's a fair amount of the protein. Each of those proteins that are enzymes that we learned about, for example, in glycolysis in a cell, if you crack a cell open they're made in larger quantities. However, all the biosynthetic enzymes, the things for making amino acids, for making purines, pyrimidines, nucleotides or vitamins, which are only needed tiny amounts, there are only little tiny bits of those enzymes. So trying to work out the biosynthetic pathways for how those things were made was much more difficult. And it was helped by genetics in the following way. I'm not going to go into any great detail, but if you imagine that there's a pathway for making histidine. We don't know what it is. But we know that the end product is histidine. We could get mutants. His-minus mutants, which I've said up there, have the property of growing only if you add histidine. So if we had just a minimal glucose plate, just some salts with glucose and we were to streak a wild type bacterium on it, it would grow just fine. But if we had a his-minus mutant on a minimal glucose plate and we streaked it out, it couldn't grow because it couldn't make histidine. If it couldn't make histidine it couldn't make proteins. But if we take the same plate, minimal glucose plus some histidine, now this same mutant will grow fine. And that's how we could tell that it was specifically broken in making histidine. So if you were to go into an undergrad lab, and I won't talk for the moment on how you would isolate those his-minus bacteria, it's not hard, you can do it in an undergrad lab, you can make a whole lot of mutants that were broken in making histidine. So one of the sort of things biochemists could do was they didn't know what the intermediates were, but let's just say there's intermediate-1, intermediate-2, intermediate-3, intermediate-4 and intermediate-5. And finally it goes to histidine. Well, if we break the gene that makes that, what will happen in that pathway is the cell will make this intermediate, this intermediate, and then this one will kind of build up. And, furthermore, if I'm able to figure out what intermediate-4 is, if I add intermediate-4 to this mutant it will grow because the defect was earlier in the pathway. And if I added intermediate-1 it won't grow. And so by sort of using very fine features of the phenotype, I mean getting in there, breaking it open and discovering when intermediate was up, or I could add an intermediate back and forth, sort of playing with the phenotype you can learn a lot about the mutants that you've isolated and so on. But genetics as -- And we'll sort of go into this just a little bit more. But genetics as a science is able to figure out whether mutations are in the same genes and work out their order without having to do any of these sort of specialized knowledge of a phenotype. It's a very general, very powerful way of doing business. And I want to show you that. I want to use a system where we can see this very clearly. And I want to introduce you to a bacterial virus. The idea of a bacterial virus, which are called bacteriophage, but it's just a bacterial virus. And what's a bacteriophage? Well, it's got a protein coat of some kind. They don't all look the same, but just some of them look like this. So this is protein and this is DNA. That's its genetic material. And it's basically a syringe. It's got a coat and it's got stuff that will let it find a bacterial host. And it's able to squirt its DNA from the bacterium into the host. So if we take, for example, an E. coli cell and this phage, which I'm drawing not to scale, it would be smaller than this relative to the bacterium. If it were to infect, it would inject its DNA into the E. coli cell. And what it's sort of done is it's put a bunch of new genetic information into a cell that's all capable of making RNA and proteins. And so it kind of reprograms the cell in the way I'm going to show you in a minute. So when this new DNA comes in, this is the E. coli DNA. And the virus kind of takes over the cell. And what it does is it makes it into a machine or a factor for making baby virus, if you will. So first it makes phage DNA. And then it also makes the proteins that self-assemble to give [its code?]. So it's kind of reprogrammed the cell. And some of the viruses that infect our cells are essentially doing the same thing. They stick their genetic material into our cells and it takes over. So unlike the retrovirus that we were talking about, this thing is not inserting its DNA into the genome of the host. It's just using the cell as a factory for making more of its own. So then these assemble so that the host has now got phage

inside it like this. And the cells then lysis, which means that they burst open. The phage, once it's done all of this, makes a special enzyme that degrades the bacterial cell wall. And it makes the cell pop open. And it releases these free phage. And if you start with one of them you might get, for example, 150 coming out of the cell when it bursts. And then each one of these is able to grab hold of another uninfected E. coli and start the cycle again. And the cycle takes usually something like 20 to 30 minutes for a bacteriophage. So it's pretty quick. One phage absorbing to a bacterium, injecting its DNA can make that 150 copies of itself in about 20 minutes, and then each of those can infect. So how would you detect something like this? Well, the trick that's done is pretty simple. You take something like ten to the eighth bacteria. Let's say you have some bacteria in a test tube and then maybe let's say ten to the two phage, and then you spread it on just a Petri plate that the bacteria can grow on. So there are many, many bacteria. So they're just going to sort of grow up and form kind of a wand that will cover the whole plate. And if you were to hold it up to the light you'd see it's sort of opaque now because the bacteria have grown up. But anywhere there was a phage to start out with, a single phage it would infect its original cell, the cell would burst open, it would release 150 phage, they'd infect the nearest 150 bacteria, they'd break open. And so what happens is the bacteria trying to grow and cover the plate, the phage are growing and eating all the bacteria. Basically at least using up all the bacteria in that area. So what you get from this are little holes in the lawn, and they're called a plaque. That's the technical term. Here. And so it's actually very easy to see how many phage you have because you just put them out on a plate. And I realize this might sound slightly fanciful. There is a textbook picture of a bacteriophage with the DNAs all packaged up in the head. It's got a sheath. It's got little things that will let it recognize a particular host. And it really does look like a syringe. And it's got stuff that squirts, basically the mechanics to squirt the DNA in. There's an electron micrograph. So that's a real one. This is not just a textbook cartoon. It's a pretty accurate depiction. This was a little thing. This is cycling, so don't get it mixed up. This only happens once. But this is basically depicting the idea that the DNA starts out in the phage and then it gets injected into the cell. And once it's in the cell the empty coat, it doesn't have anything else to do in this story, but the DNA takes over. And then you get, after a little while, when you've made these progeny phage the cell breaks open. And if you assay them, this is just a picture one of my post-docs made for me. I think you can see here's a lawn. You can sort of see how it's opaque and you can see those little holes. There were probably a thousand or so phage that were put on this plate. There are several hundred anyway. And you can just count the number of holes and then you know how many phage you've got. OK? So that's this system. Now, what I'd like to now consider is how we could use genetics to try and study the essential functions of that bacteriophage. How does it replicate? How does it make its coat? I want to study the things that are needed for it to be a phage, not something that's dispensable. I want to know essential functions. So that means I would have to use conditional mutants of some type so that I could study it because otherwise if I broke something that was critical for the phage's life cycle I'd never get any phage and I couldn't do any genetics. So what I would do is I would look for temperature sensitive mutants. The phage. So they'll form plaques, let's say plaques at 30 degrees. No plaques at 37 degrees. And you go through, and you could be laborious or clever depending on how you set this up. But we could make a bunch of mutants. And let's call T1, T2, just give them names like that as I isolate them. Now, what I want to show you are two absolutely standard and critical genetic ways of analyzing these mutants. They all look the same. Their phenotype is they form plaques at 30 degrees. They don't form plaques at 37 degrees. They could be affecting one gene. We could have mutants in 50 genes. I don't know starting out. All I know is I've got mutants that are temperature sensitive. So genetics gives you a couple of ways of going at that that will tell you not only are they in the same gene or in the different gene. But it also will tell you something about their physical relationship along the DNA. And this is without knowing anything about what they do. So let me show you how that works. So the first genetic operation is called a complementation test. And the thing that I hope will strike you about particularly these bacteriophage things is these are so simple. I've already told you basically all the techniques we're going to be using. We're going to be taking phage, we're going to be mixing it with bacteria, we're going to be putting them on a plate and we're going to be counting plaques. But we're not going to do anything else and we're going to learn stuff about whether mutations are in the same gene or different genes, between different mutants and something about the order of the genes on the chromosome. And that's the point I'm trying to drive home right now. So here's the first idea. Let's add the T1 mutant plus the T2 mutant to some bacteria. So we'll put them both into the same test tube. And what we want is enough phage -- So every bacterium gets both. So we want to take, if we take T1 by itself it will grow plaques at 30 degrees, not at 37. T2 plaques at 30 degrees, no plaques at 37. Now we're going to take some bacteria and put enough that both of them will get into the same thing. Now, of course, those bacteria are doomed because if anything happens, because they've got things inside them. So what we do now is then we add a whole bunch of what we would call indicator bacteria. These are ones that haven't been. And we'll put lots of those in, and we'll put many fewer of these and we'll mix these together. And then we'll plate them out under two different conditions. Let's try it at 30 degrees. Well, in that case, I think we can figure out what would happen. Both T1 and T2 can form plaque, so you'd expect to see plaques. And you would indeed see that if you did this experiment. Now, the other thing, though, this is where it gets interesting, is what would happen if we plated them at 37 degrees? Well, on their own neither of them can form a plaque. But we've engineered it so that there are two within each bacterium. So there are two possible outcomes, and let's think what they could be. We could either have mutations in the same gene that some particular gene, let's say a gene required for making the major protein in the coat, that mutant one is affected in that gene and mutant two is affected in the same protein. So both of them got a broken coat protein. And so inside this bacterium, when the phage are trying to grow, what you've got is this one gene. And this is the T1 mutant and this is the T2 mutant. And maybe this one has got a mutation somewhere and the other one has got it somewhere else in the gene, but this gene is not functional. So we wouldn't get any plaques. But what if they were in different genes? Let's say one of the mutants was altered in the gene for a coat protein and the other one was altered in a gene that was necessary for replicating the phage DNA. So it's a situation like this. And let's say this is gene A and

that's gene B. What do you think could happen now? Get plaques? Wouldn't get plaques? Yeah? I see a lot of nodding. We'd get plaques because this one has a good gene A, so it would make let's say the coat protein. This one has a good gene B so it could make the DNA polymerase for copying the phage DNA. And things would be fine. And so by this very, very simple test, we could take a whole lot of TS mutants and we could go through in a pare-wise fashion, and we could say oh, I see, number one, number seven and number 54 apparently affect all mutations affecting one gene, and we could put them into categories by doing this. And we haven't done anything other than mix phage and bacteria and look to see whether they're plaques or not. I mean this is very different than what a biochemist does. It's a different kind of thinking. And yet it's enormously powerful. So this procedure, which is one of the workhorses of a geneticist, is known as a complementation test. And depending on the organism it takes a whole bunch of different sort of technical forms. But that's the principle of the thing, that if you have one good copy of the gene and a broken one then you can survive in most cases because usually the one will be enough to get you through here. So that's one of the things that geneticists do. And see how powerful it is for something that's a very simple manipulation. But let me now tell you the other kind of test that geneticists do. It's a slightly different principle but just as powerful and gives you a different kind of information. This is known as a recombination test. So what we're going to do in this case now is we're going to allow two mutants to grow together in the same cell. Let's say two mutants. And we'll use T1 and T2 again to grow together in the same cell. Now, for example, remember up here we mixed both of them in this so they both were inside this bacteria and then when we plated it up, up here at 30 degrees we got plaques? We would have gotten plaques from every single infected bacteria. Well, those plaques probably have ten to the ninth phage in them. And those were phage that grew together under permissive conditions. OK? So let's take those, re-suspend them -- -- and bring them over here. Now, we'll add them to some bacteria. So bacteria in here. But we're going to do it under different conditions now. We don't want complementation taking place, so we'll make sure that we have less than one phage -- -- per bacterium. So we cannot do this complementation thing anymore. We just got rid of it by using a different ratio of phage to the bacteria. So now what I want to do, I'm going to plate this out at 37 degrees. We could add an indicator if we wanted again. Well, that may seem like a stupid experiment in the sense that T1 wouldn't grow by itself and T2 wouldn't grow by itself at 37 degrees. And all I've done is let them grow together. So if all we had in that population was what we started with there wouldn't be any plaques. But if you do that what you will find is you'll find some rare plaques. And these are what are known as recombinants. And let me now just give you a sense of how these recombinants arise. So let's imagine that this is the mutation and here's our T1 bacterium. Here's its DNA. And let's say there's a function here that it's wild type for and it's got a mutation down here. The other one, T2 has got the mutation here but it's wild type for there. That's the kind of thing we were talking about, a gene A and a gene B, although it's more general than that. But what happens when these things are growing together is under rare circumstances this interesting thing happens. Since this is the two strands of DNA and since these phages are almost identical and this piece of DNA is the same as that piece of DNA. And if a break every happened in one of these strands, what can happen is it can invade the other strand and displace it. And this strand can do a little switcheroo and come over here like this. So now we've got the plus and the plus and the minus and the minus. So this is an intermediate but these things happen. You get little breaks in DNA from DNA damage and other things. And the other piece of DNA can go off and pioneer and find another molecule. And then there are enzymes that resolve this. And if we cut it right here, we'd just go back to what we had before. But if you cut this strand here, and you may have to sit down and draw this out because this, for me, was not intuitive when I was learning this thing. You'll see what happens now. The end of this strand will go over and join to there. The end of this strand will join and go here. And what you will get out of that is one phage that's got both of the pluses and one of the phage that's got both of the minuses. And I'll call this sort of T1, T2 to indicate that it has got both of the mutations. So this would be what we could detect over there. That would be one of the rare wild types. And I'll assert to you that this guy here has two mutations in it instead of one. What's the phenotype of the one that has both mutations? It cannot grow. It can make plaques at 30 degrees. It cannot make plaques at 37. If I got up and said I'm sure this phage has two mutations. It has both T1 and T2 mutations in it. And I say prove it to me. Anybody see how you could do it given what I've already told you? Exactly. If I tried it in a complementation test with T1 it wouldn't compliment because both would be broken in T1. If I tried to complement T2 it wouldn't work because it's broken in T2. And just by using these tiny little simple manipulations that I've told you, I can even see that something that has got a double mutant we could sort it out at the bench. You could walk in and tell me the next morning what the result was. In fact, phage grows so fast you might even be able to do it in the morning and tell me before you went home at night. So this is, as I say, called a recombination test. And it's giving you a different kind of information, but it's an extraordinarily powerful technique for another reason, is that the recombination frequency can be measured. We'll call RF. And this is the recombinants over the total. So, in our case, this would be the number of wild type, plus the number of these double mutants over the total of T1 plus the total of T2, which would be the dominant members of this population. Because this is a relatively rare event, plus the wild type, plus the T1, T2. And the thing that is so useful is the probability of, geneticists call this a crossover, this kind of crossover happening is proportional to the distance between these things. If they're a very long distance apart there's a lot of DNA, and the chances of it happening are much higher than if the two mutations are very close together. It makes sense just from First Principles. So the recombination frequency varies as the distance -- -- between the mutations. So let's imagine that I do a cross like this. We'd grow T1 plus T2 and we measure the recombination frequency, and I find that it's 4%. 4% of the plaques are mutants out of everything that's in there. And let me try another one. I'll take T2 and I'll cross it with the next mutant phage of isolated T3. And let's say, in this case, the recombination frequency is 5%. Because I know they're different distances, but if you think about it you'll realize there are two kinds of maps we can draw. These data are compatible with the gene order being one, two, with this being a distance that corresponds to a 4% recombination frequency, and three with this being the distance that corresponds to the 5%. But it's also compatible. isn't it. with this where we have one.

two and three here. There's the 4%. And from there to there is the 5%. Yeah, that's right. Two to three. One to two is 4% and two to three is 5% in both of these. One to two is 4%. What did I do wrong? Two. I guess I've got to do it this way, two, one. Is that going to do it? What did I do with my example here? Yeah. Hang on a second. OK. Let's go back to where I was and see if I can reconstruct this. OK. There's one, and two is 4%, and we want two to three 5%. OK, right. So it's going to be here. Here we go. Two to three. Excuse me. It's got to be farther over here. There's three. There's the 5%. OK. Now we've got it, once I change these numbers. The 4%, the 5%, the 4%, the 5%. OK. I thought I was over this cold. I guess I'm not quite over this cold. OK. Could you devise an experiment that would distinguish between those maps? Yeah? Yeah. You've got it. Take T1 and T3. One of those you'll get a 9%. One of them you'll get a 1%. Isn't that amazing? I mean it's so simple, but I think you can perhaps see some of the power of genetics here. All we've done is we've got things that we know can form plaques at 30 and not at 37. We haven't done anything other than mix them with bacteria and then count plaques. And we've been able to make inferences in a living organism about mutations in genes that are absolutely essential for that organism to grow. So what phage geneticists were able to do over the years then was they would make a map of all of these different genes down to the phage genome and then the biochemist would go in and work it out. And people studying mice would make maps of genes along mice. And it's only been in the last handful of years we've been able to go in and sequence the DNA and find out where all of these things are. So I will see you guys after. I hope you have a wonderful spring break. And I think you'll really enjoy hearing Penny. And we'll start in on diploid genetics and Mendel and stuff as soon as we get back. OK? So see you in a little bit.