ERIC LANDER: Good morning. Good morning. So last time, we ran into a problem. We had Mendel, my hero Mendel, this MIT-like mathematical, physical monk, had developed this gorgeous theory of particles of inheritance.

He didn't use the word "gene" yet. Gene doesn't get invented for much longer. For every trait, you had two such particles. You gave one to your offspring. Each of the parents gave one to their offspring. And that's how each offspring gets two of them.

That choice of which of the two alleles to transmit to your offspring is a random draw. And that explains beautifully, for example, the 3-to-1 segregation pattern that Mendel saw-- gorgeous. We put that model, which was an ex post facto model, a model made after the data were available, to a test. You guys insisted we had to test it before you would publish it.

And it holds up pretty well, making pretty surprising predictions that you would otherwise not have ever expected. Like amongst that 3 to 1, the threes are not all the same. Some of those round peas were homozygous for the big-R allele. And in selfing, they'll never produce wrinkled peas.

But 2/3 of those round peas were heterozygous. And when you self them, you get 1/4 wrinkled peas. So the wacky prediction that 1/3 of the rounds will give rise to no wrinkleds, and 2/3 of the rounds will give rise to 1/4 wrinkleds is a surprising prediction and, therefore, something that bears stating as a scientific prediction. And so all those kind of predictions can emerge.

Then we turned to the question of two factors. I'm practicing using our words here-- homozygous, heterozygous, alleles, et cetera. We now turn to two traits, two phenotypes. We have two phenotypes segregating. We had round and wrinkled, and we had green and yellow.

And Mendel determined, rather brilliantly, that they segregated, were transmitted independently of each other. There was no correlation between which alleles you got at round and which alleles you got at wrinkled. That was pretty cool. Let's just go

over that, because there's a real tension to be resolved because we have Mendel's second law versus the chromosome theory.

So Mendel's second law-- let's just go back to it-- in the F0 generation, we had our round green peas, genotype big R, big R, big G, big G. We had our wrinkled yellow peas, genotype little r, little r, little g, little g. We cross them together, we get F1 double heterozygotes, big R, little r, big G, little g.

We then perform a back cross or test cross to the doubly homozygous parent with the two recessive phenotypes-- we're practicing our words here. And what do we get? Well, we get certain options. As we said, the gametes that emerge from this parent on the left could be of the following types.

The gametes from the parent on the right are all those alleles-- Are those the recessive alleles? No. You'll forget this over time, but just to make me comfortable-- they're actually the alleles associated with the recessive phenotype that we're talking about. Because you know-- but will forget, I assure you, because all my colleagues in the biology department have forgotten-- that they could also control multiple other phenotypes, some of which could be dominant. But that's OK. I forgive you in advance that you'll call them recessive alleles.

Anyway, you get this. And then these should occur at equal frequency of one to one to one to one. All right.

Now, we had the chromosome theory. The chromosome theory, the observation of the choreography of chromosomes during meiosis-- chromosomes in meiosis-- looks like this. We have chromosomes lining up in pairs. Now, I'm not drawing it terribly well.

But the two members of each pair are of the same size. But this pair could be bigger, and this pair could be smaller. It looks like they're really finding each other. These chromosomes are visibly different in shape. And so maybe I'll make this guy a little longer, just to indicate that, that it actually knows.

Now an explanation, we said, for how Mendel's second law of independent

assortment of two different phenotypes could occur is that, for example - round. It could be that the gene for roundness is located on chromosome number one. These are two copies of chromosome one that have been duplicated, each of which has the big-R allele. Next to it, two copies of chromosome one that have been duplicated, each of them having the little-r allele.

Over here on chromosome number two. let's say, lies the gene for green or yellow. And in this picture here, the big G's are on the two copies of this chromosome number two that are here. The little g's are on the two copies of chromosome number two that are here. When the cell undergoes meiosis one, we get to the situation where we have big G, big G; little g, little g. We've got big R, big R; little r, little r.

Could it have been the case that the big G was on the right and the little g was on the left? Yeah, of course. It's totally independent which way. I happened to draw it this way. But with probability 50%, it's the other way. That's why they're independent of each other.

And then when it undergoes meiosis two that looks just like mitosis, we end up with our four gametes here. Sorry, our four gametes with big G, big G; little g, little g; big R, big R; little r, little r. And that accounts for the big G, big G; big R, big G; little r, little g gametes.

And then when they went the other way, the organism would make a set of gametes that had the other combination of big R's with little g's. So that's perfectly fine. And because the second chromosome is independently ordered compared to the first chromosome-- when they line up at the midline, they don't really care which way they are-- that'll account for one to one to one to one. It's so straightforward.

But what happens if, instead, both the roundness gene and the greenness gene live on the same chromosome? We'll have little r, little r there. We'll have big G, big G here; little g, little g.

And then when they split, we'll end up with-- now, chromosome two has nothing that

we care about on it. It has a lot of genes. In fact, there could be chromosomes three, four, five. I'm just not drawing them. But we're really going to focus on chromosome number one here.

And you'll notice that here on chromosome one, the big G's or the little g's are coupled, physically coupled to each other, the bigs with the bigs and the smalls with the smalls. So now the kind of gametes that can emerge from this will only be big R, big G type or-- big G type-- or they'll be little r, little g type. We can't get the reverse combination. We can't get bigs and littles.

So this, because these are independent of each other, will get us one to one to one to one. These are the big, little, and then these other combinations like that. This will get us only, if we look at the big R, big G; little r, little g; big R, little g; little r, big G; will get us one to one to zero to zero.

Let's give a name to this type, the big R's and the little g's. Let's call that a recombinant type, OK? I'm just going to use that word for the moment. The recombinant types-- the bigs and the littles, and littles and the bigs-- we're not going to see any of them.

This is a very strong difference between Mendel's second law and the chromosome theory. If the chromosome theory is right, and these chromosomes are physical entities that have integrity, they can't both be right. So that's a great thing in science, when you have two different models and they can't both be right, because you learn things then. You can test them.

Now, Mendel tested this without actually knowing the chromosome theory. And he always got one to one to one to one for the seven traits he looked at. Was he just lucky that they happened to lie on the seven different chromosomes? Or is there some problem with this chromosome theory, or what?

It took a while. And then, of course, everybody forgot about Mendel from 1865 until the year 1900. In the year 1900, people begin to rediscover Mendel. Cytology has come along. In January of the year 1900, plant breeders start rediscovering

Mendel.

Sorry. Thank you. I see already. Thank you. I could tell by the look on your face that something must be wrong there. Good.

Plant breeders start rediscovering Mendel. And in January of 1900, three different groups say, you know, we found these laws. And they're just like Mendel's laws-- which now everybody starts paying attention to. But plant breeding-- and people tried to do mice, and people tried to do rats.

What turned out to be the winner, the place to really study genetics, was the fruit fly. Thomas Hunt Morgan at Columbia University decided, after he was really frustrated wasting years breeding mice and rats that just took too long, around, oh, I don't know, 1906 or something like that, began to start breeding fruit flies. Fruit flies are the teeny little flies that when you open a banana or fruits or other kinds of things, you'll see them. And studying Drosophila gave us the answer to this question of what the problem is, how can it be that either it's independent or totally dependent?

So we're going to talk about Drosophila melanogaster, the fruit fly, and the discovery of recombination. So now, Morgan-- I'm going to now start using fruit fly notation to give you some practice with fruit fly notation. We're not going to use big G's and little g's.

They like to refer to the normal allele and the mutant allele. The normally allele is plus. The mutant allele gets some kind of a name.

And so he had a female fly that was normal and normal at two different loci, two different genes. I haven't told you what the genes are. And the male fly had a black-colored body-- black across its whole body-- and its wings were shriveled and, therefore, called vestigial.

So the phenotype here is black and vestigial, black body and vestigial wings. And this wasn't the normal appearance of a fly. So he took these females by these males, and he crossed them together. And he got an F1. And the F1 was black over vestigial, plus over plus.

And what was their phenotype? Were they normal appearance, which is kind of a sandy-colored body and normal wings? Or were they this all black body and vestigial wings?

Turns out they were normal. From that, what do we infer about these two traits, black body and vestigial wings? They're recessive traits.

So here, the phenotype was normal appearance. So then he crosses them back, doing a test cross. Let's say he'll take males here and females here, but it actually works either way. And what he does is just like we did there. He could get gametes that were black, vestigial. He could get gametes that were black, plus; plus, vestigial; or plus, plus. Those are the four possibilities that come out.

So when he does it-- let's keep score. I'm going to write them now-- plus, plus; black, vestigial. And from the other parent, you got this-- black, plus; black, vestigial; plus, vestigial; black, vestigial. Those are the four possibilities. And if this was just like Mendel's traits, it would be one to one to one to one.

These were the parental types that went in. Plus, plus went in. And black, vestigial went in. Those were the combinations of traits here. These were new combinations. What he observed-- Let's see. If Mendel's right, it'll be one to one to one to one . If the chromosome theory's right, it'll be one to one to zero to zero. And who was right?

STUDENT: No one.

ERIC LANDER: No one. The answer was 965 to 944 to 206 to 185. Neither model was right. Neither model's right.

The new combinations, the recombinant combinations, the non-parental combinations-- we can call these recombinant combinations. These were recombinant. They recombined in some way. They were a new combination, or they were the non-parental types. We use both of those words frequently-- were neither equal nor were they completely absent. They occurred, but at a lower frequency.

What was the frequency? Well, we could just add it up. The frequency of recombinant types, of new types of were different than the parents', is 17%. What's going on?

Now, maybe this is some magic number like the 3-to-1 ratio. And you should look at that 17% and say, ah, this is some constant of the universe, that when you put in traits you get 17% percent of recombinant types. But it takes a little judgment to say, I don't think so.

And he actually tried other traits. And sometimes he got one to one to one to one. But very often he got some funny number-- 6%, 28%, 1%. There was some funny business going on.

What's going on? Recombination. That's what's going on is there is recombination occurring. What do we mean by recombination?

Recombination is very important stuff, by the way. At some point, I will tell you that understanding recombination was actually the origin of the Human Genome Project. And it traces back to a good MIT story. But that will be for a little later in the semester.

So what do we think's happening here? I'm now going to draw a close-up of that chromosome. And here's another chromosome, the other pair.

And what we think here might be happening is that you might have plus and black; black and plus-- oh, sorry. Black, right? Black, black; plus, plus. This would be plus. This would be plus. This is the normal chromosome. This is plus.

So this chromosome here carries the pluses. This guy here carries those alleles black and vestigial. Well, what happens, the idea was that somehow these chromosomes exchanged material here. And the chromosome that was plus, plus; plus, plus now somehow acquires that bit, and this chromosome somehow gets that bit.

And we end up instead with a picture like this, where some of this came from here.

And those two loci, black and vestigial, were separated from each other such that the black allele moves over to that chromosome. Is that clear? That's the notion.

Why did they think this was true? Well, it turns out that in fruit flies, you can actually look at eggs under the microscope. And you can look at the chromosomes. And if you take if you take the cells, and you take a cover slip, and you squash it down with your finger, you can actually see chromosomes lying right over each other, making little crosses like I drew there, up there, called chiasmata, which means crosses.

And so people said, see, in the microscope, you can see they're lying on top of each other. Are you impressed by that piece of evidence? No. You took the cover slip, you squished it down with your fingers. So they're lying on top of each other? Big deal.

I'm not going to be impressed. And calling it chiasmata doesn't make me any more impressed, right? Although it's always good to call things Greek names, because people think they're more sophisticated if you call them Greek names.

But this was the notion people had. And the frequency 17% would indicate how often these crossovers occur. But if I were in the situation we talked about with Mendel, and I wrote this up, and I said, see, it's 17% sometimes, 6% sometimes, 28% sometimes; and when I look in the microscope, they lie on top of each other; therefore, it's recombination-- There were actually other ideas floating around too. Maybe it has something to do with developmental biology. It was a puzzle.

When you have a really deep puzzle, the most important thing in science is to find a young person. Because young people come without prejudice. They say, let me just look at all of this. Stand back. I don't come with any prejudice.

So at MIT, what is the solution when you have a problem like this? A UROP. You want a UROP. So even 1911, that was the solution at Columbia. Thomas Hunt Morgan got a UROP.

I'm serious. He was called Alfred Sturtevant. Alfred Sturtevant was a sophomore at Columbia. Everybody else was busy finding this recombination data, how often this

recombined with this, this with this, this with this.

Sturtevant-- he's a sophomore-- he says, god, this stuff's interesting! Professor Morgan, could I have all the data and try to look at it? Sturtevant took it home and actually pulled an all-nighter, blew off his homework-- it actually says so in his autobiography. He says, I blew off my homework and pulled an all-nighter-- essentially in those words, he says. "To the detriment of my undergraduate homework" is the way he puts it.

But in any case, genetic maps and Sturtevant's all-nighter. What Sturtevant did was the following. He said, how are we going to prove the chromosome theory? I like this idea that recombination is about distance on the chromosomes. I like the concept that maybe 17% is how often these things recombine.

Why would things only recombine 1% of the time? What would that mean? They've got to be pretty close together so that a crossover between them happens not so often. And what if things were far apart? Well, it could be more likely.

So he likes the idea that recombination frequency means distance. But how are you going to prove that? It could mean a zillion other things. It could mean biochemical pathways, developmental biology. How are you going to prove that recombination frequency means distance?

You've got to make predictions, right? The only to do it would be to make predictions. So Sturtevant takes the data, and he starts making predictions.

I think, says Sturtevant, these things are alleles living at genes with locations on the chromosome-- black, vestigial. How often do black and vestigial recombine with each other? What is the frequency of recombinant non-parental types? 17%.

So Sturtevant goes through the data and he says, what about other crosses people did in the lab? Well, it turns out people did crosses with another mutant that produces a funny eye color called cinnabar, Cn. So cinnabar.

It turns out that the recombination frequency between cinnabar to vestigial was 8%.

Vestigial, cinnabar, 8% recombination. If this chromosome business is right, where should I put cinnabar?

Sorry? Where do you want it?

STUDENT: [INAUDIBLE].

ERIC LANDER: Somewhere in between. You'd like me to put that there.

STUDENT: On the other side.

ERIC LANDER: Oh, OK. Wait a second. On the other side.

OK, which is it? How many vote for the left? How many vote for the right? How many conscientious abstainers are there?

Do we know?

STUDENT: No.

ERIC LANDER: No. There are two possibilities. It could be 8% this way, or it could be 8% that way. How are we going to know? Yes?

STUDENT: [INAUDIBLE].

ERIC LANDER: Black. What if we knew the answer between cinnabar and black? That would constrain the problem. Can you give me two predictions for what the answer might be?

STUDENT: Either 9% or--

ERIC LANDER: Either 9% or 25%. So now we have a prediction. We don't know where cinnabar is. But the answer could be that black to vestigial should either be about 9%-- that's what that would be here-- or about 25%.

The answer? About 9%. That's what Sturtevant found. That's a prediction and kind of cool. And you can imagine taking the data home and it's probably 9 o'clock, and you've now realized, wow, freaky, it's 9%.

So then he looked at the mutation "lobe." Lobe was another mutation. The lobe mutation showed 5% recombination from vestigial. Where should we put it? Left or right, well, let's make some predictions.

Suppose it's over here. Will it be very close to cinnabar? Will it be closer to black? But what if it was over here? Well, it would be further.

So let's put lobe in. And suppose we know that lob is 5%. Then what's the prediction for black to lobe? 22%. Answer, according to the notebooks, 21%, pretty close.

You'd like it to be exactly 22. But life doesn't always turn out that way. 21's pretty close to 22.

What other predictions could we make? If this is cinnabar here, could you give me a prediction for cinnabar to lobe? 13%. Yep, that works.

Curved wing. Recombination distance from lobed, 3%. So now you have some predictions.

You have this prediction here. You predict 8%. Answer, about 8%. Over here you predict 16%. Answer, about 16%. Bingo.

Sturtivant says, if these genes-- we call them loci, often. I'll use the word locus synonymous with gene. Locus means a place. And geneticists think about genes as a place on a chromosome.

If these loci-- the plural of locus-- if these loci were really arrayed along a linear structure, then it would have to be the case that they would have certain additive relationships between them. And the chance that they would have these additive relationships if they weren't part of a line is pretty implausible. That's a real prediction, a very remarkable prediction. And it holds up with the data.

Sturtivant pulls the all-nighter. By the time the sun comes up at Columbia University-- this is Morningside Heights-- he's got the whole thing worked out. Yes, this chromosome theory must be right. It fits beautifully all of these data. Pretty cool.

You are all authorized to blow off your homework anytime you make a discovery like that. [LAUGHTER] That's a course rule. Any homework will be forgiven for discoveries of that magnitude. All right. Tell your TAs.

So now, what does it really tell us? It tells us that if genes are very close together, the recombination frequency, R recombination frequency, or RF, might be very little. They could be as low as almost zero, which means you never see a recombinant because they're right next to each other.

Or it could be that they're further apart. It might be 1%. It might be 10%. It could keep growing. It might be 30%.

Suppose it's way, way, way, way, way far away. What's the largest it ever gets to be? Well, if they were on different chromosomes-- suppose there were totally independent segregation, different chromosomes-- what would they be?

STUDENT: [INAUDIBLE]

ERIC LANDER: No, it's 50% because remember, one to one to one to one says that the recombinant-- well, actually, this is an interesting question. Let's come to that 100. On different chromosomes, one to one to one to one means that it's 50%, half of them are recombinant types. It turns out on the same chromosome, as you get farther and farther and farther, you might say there's going to be 100% chance of a crossover. And you might say the recombination frequency could keep growing past 50%.

It turns out it doesn't. The reason is that multiple crossovers can happen. So mathematical interjection here, if here's my gene and here's my gene, there could be one crossover. There could be two crossovers. There could be three crossovers.

And so in fact, it turns out, as you get very far away, you have to start paying attention to the probabilities of double crossovers and triple crossovers. And so it turns out that it's a Poisson process of the number of crossovers that occurs, give or take. And you see a recombination if there's an odd number. And for that reason, it

never gets above 50%.

So as the distance gets further and further and further, it goes from zero to 50, which is the same number you get for separate chromosomes. Otherwise, you might think that if there was just one crossover, it gets to 100% probability of recombination. But it never does, because there are doubles.

And you can actually observe, if you make a cross that has three different genes segregating in it, you can actually see the double crossover type. So you can see very nicely that if you make a cross involving black, cinnabar, and vestigial over plus, plus, plus, you can see that cinnabar is right in the middle because this recombination happens at a pretty good frequency.

This recombination happens at a good frequency, giving you plus, cinnabar, vestigial or black, plus, plus. Sometimes you will get that recombination. You'll get out gametes that are black, plus, vestigial, but at a much lower frequency because they take two crossovers.

What will be the probability of seeing a black, cinnabar, vestigial? Well, we said that this one was about 9%, this one was about 8%. What's the product of a 9% chance and an 8% chance? It's about a 1% chance, a little less than a 1% chance. That how frequently you see black, plus, vestigial.

You can even predict the probability of a double crossover event by multiplying the two events that have to happen. So your bottom-line rules here is that because of these double crossovers our recombination frequency can go from zero to about 50%. This is independent assortment. And it either occurs if you're on different chromosomes or, if you're very far away on the same chromosome, they behave as if they're independent of each other.

Any questions about any of that? Yes?

STUDENT: Why didn't Mendel ever see recombination?

ERIC LANDER: Why didn't Mendel ever see recombination? It turns out that with seven

chromosomes, and they're biggish in length, he never actually ran into two loci that were close enough, to notice it. That's why.

Flies actually only have three major chromosomes. There's a fourth, but it's a puny little thing. And because they were much more intensively collecting mutations in Morgan's fly room, they began having a lot of them, and they had to bump into recombination pretty early. Mendel simply didn't have enough that were close enough.

Think about what would have happened if just by chance, somebody were selling a strain of peas in the market which had a mutation in a locus that had 10% recombination distance, and it screwed up Mendel's law of independent assortment of two loci? Mendel might not have published the paper. Sometimes in science it's actually valuable to first get the oversimplification out , like his second law, and then deal with the complexity that sits on top of the oversimplification.

It's kind of lucky that Mendel didn't have enough of them to be bothered, in the first paper. So in fact, all of Mendel's seven loci have now been mapped. Most have been cloned molecularly. And so we actually know where they are, et cetera. It's a really good question. That's sort of why he didn't.