

Discovering Genes Associated with Diseases and Traits in Dogs

[00:00:00.88] [MUSIC PLAYING]

[00:00:21.40] Hello. My name is Dr. Elinor Karlsson and I work at the Broad Institute of MIT at Harvard. And today I'm going to give you a lesson that tells you about some of my research collaborators. I work with dog owners from all around the world who donate DNA samples from their dogs. And then I look in these DNA samples and try to find the genes that make dogs look the way they look, make dogs act the way they act, and also find the genes that make dogs sick.

[00:00:46.05] And so why are we doing this in dogs? Well, dogs are an absolutely amazing species because they've lived with us for thousands and thousands of years. And you can see here that these Golden Retrievers are playing very happily with these people. And that's because for these thousands of years we've kind of depended on each other. Humans have loved dogs, they've fed the dogs, and in return dogs do all sorts of useful behaviors for us.

[00:01:07.70] The Golden Retrievers are really wonderful at finding things, and then bringing them back, and that's why they're called retrievers. And then there's other breeds that herd sheep, there's breeds that are really good hunters or helping hunters, and then there's also breeds that are very protective and they'll protect people's homes. Here I'm showing you four different dog breeds and each of these dog breeds has been bred to do a very specific behavior. And I'd like you as a class to now look at these breeds and figure out which of them you think matches up to each of these behaviors. And then we'll come back and discuss each of these breeds.

[00:01:50.54] Now that you've had a chance to look at these four breeds and discuss what you think they might do, let's go through them each one by one and discuss what they're called and what they've actually been bred for.

[00:01:59.81] So the first breed we're going to discuss you can see on this slide is actually a tracking breed, and it's called the Dachshund. It has these funny little legs and that was actually partly what they were bred for. They have big, long noses that have a great sense of smell so they can track down animals. And because they have short stumpy legs, they can actually fit into burrows underneath the ground so they can track things like badgers that live in those burrows.

[00:02:23.18] The next breed is our racing breed, and this is called the Greyhound. And you can tell that it's a racing breed just by looking at it. You can see how long the legs are, you can see how big its chest is, so that it can fill it up with lots and lots of air. And it turns out these are actually the very fastest dog breeds, they can run over 60 kilometers an hour, or 40 miles an hour.

[00:02:43.98] And then the next breed we're going to look at is the herding breed, that's the Border Collie. These dogs are thought to be the very, very smartest of dogs. They work with farmers in the field in these huge fields, and huge flocks of sheep. And the farmers give them lots of complicated instructions, telling them what they should do. And there's actually been a Border Collie they've tested in Germany that knows over 300 different words, which is more than a 2-year old child does.

[00:03:10.11] And the last breed is our companionship breed, and these are the really tiny breeds. This breed is called the Pekingese. It weighs as little as, I think, about seven pounds. And you can see that was actually bred to be that small, it was supposed to sit on people's laps. The woman in this painting here, that I'm showing on this slide, has a Pekingese sitting in her lap.

[00:03:29.57] And so these are these four different breeds that were bred to do very different things, and you can tell by looking at them that they were bred to do these very different things. But where did all of these dog breeds come from? Well, people created them by breeding dogs together that had something that they wanted. Now when you look at, say, these four breeds again, you see they're all very different from one another. But if you wanted to create a dog breed, what would you do? So looking at this set of 20 dogs that I'm showing you here, if you wanted to create a breed of all large brown dogs, which ones would you choose? So each of you should discuss now which dogs out of here you would pick to make your breed and then we'll come back and see what happens when we do that.

[00:04:17.55] Hello. Hopefully you all had fun creating your large brown dog breed. And I'm guessing that the dogs that you would have chosen off of that worksheet to make a large brown dog breed would have been the large brown dogs. And this is actually exactly how people have created dog breeds. If they wanted a dog breed that had very short legs, they would choose all of the dogs that had very short legs and breed them together to make that breed. And if they wanted a breed that was really, really fast, they would find the fastest dogs and breed them together to make that breed. And this is how we've gotten over 400 different dog breeds around the world today.

[00:04:50.75] But there's a problem with this. What would happen if out of the large brown dogs that you used to make your breed, one of them had a version of a gene that was involved in eye-sight development that meant that it went blind at a fairly early age? Well, it would then pass that gene on to its children, and then those children would pass that version of that gene on to their children. And eventually you would end up with a dog breed that was large and it was brown, but a lot of the dogs in that breed would have a version of that eye-sight gene that meant that they went blind at a fairly early age. And this is a problem that a lot of dog breeds have. It's not just blindness, there's breeds that have problems with a lot of other diseases. And now as a class, I'd like you to brainstorm a list of what kinds of diseases you think dogs are getting and then we'll come back and discuss what they're actually seeing at veterinary hospitals when dogs come in and they're sick.

[00:05:52.68] Now that you've come up with a list of diseases you think dogs might be getting, let's take a look at the list of the diseases that are actually the top diseases that are seen by veterinarians in dogs. Looking at this list, you'll notice that a lot of these diseases are the diseases that people get, actually all of them are diseases that people get. So things like heart disease, arthritis, epilepsy, cancer. You probably know somebody who's had one of these diseases. And you might also know that doctors often have a very hard time treating people that are sick with these diseases. We hope that by finding the genes that are causing these diseases we can come up with better ways to treat them and even cure the people that are sick.

[00:06:31.79] So why are we doing this in dogs rather than in people? Well, one is because dogs are getting all the same diseases as people, that makes them very good models for the human disease. And the other reason we're using dogs is because it turns that the dog breeds, because of the way all the dogs in the breed are very similar to each other, are actually really good for finding disease genes. Let me show you what I mean.

[00:06:55.04] So if we look at these two dogs, I've got one dog here that's sick and one dog that's healthy. The dog that's sick has the red circle on it. And if you compare these two dogs, what we're looking for is a gene that looks different in the sick dog than it does in the healthy dog. But what we actually do is we don't just compare two dogs to one another, we compare 100 sick dogs to 100 healthy dogs. We say, which gene looks different in the sick dogs than it does in any of the healthy dogs?

[00:07:20.48] And this is why this works so much better in dogs than it does in people. We could do this kind of study in people. We could take 100 people that were sick with diabetes and compare them to 100 healthy people, and some of the genes that were different might have something to do with diabetes. But there would also be genes for all the other differences that you see in people. There's genes for differences in height, genes for differences in hair color, and lots of other differences that you would see if you compared a group of people. But if you do your gene mapping in a dog breed, all the dogs in the breed are

so similar to one another that most of the genes you find that are different between the sick dogs and the healthy dogs are going to be genes that are actually related to making those dogs sick.

[00:07:59.76] So what are we actually going to do here? What we need to do is look at all of the DNA that's in the dogs and find the genes that are different. So if we look at this dog here, every single one of the cells in this dog's body has a whole lot of DNA in it. And this DNA is basically the instructions for making a dog. Here's a picture of a cell and if we zoom in on that picture of a cell, you can see that this blue circle in the middle is the nucleus, which is where all of the DNA lives. And now if we zoom in even more on that nucleus and look at what's inside it, you can see that inside every single one of those dogs cells inside the nucleus, there's 78 chromosomes. There's 39 chromosomes that the dog got from its mother and 39 chromosomes that the dog got from its father.

[00:08:40.93] And what are these chromosomes? Well, every one of these chromosomes is just a long string of A's, C's, T's, and G's, like I'm showing you here. And that just goes on and on and on. And for the most part, if you look at two dogs in a breed, most of these letters are going to be exactly the same. But every once in a while you're going to find differences. So here I'm showing you two Golden Retrievers, and for the most part their DNA is exactly the same. They both start with A, A, T, T, G, and go on much the same as one another. But every once in a while you see a difference. And so here you've got an A in the top dog, and you have a C in the bottom dog, and here you have a G in the top dog, and a T in the bottom dog, and we call these differences SNPs. So now I want you to discuss with your neighbor how many letters of DNA you think we're going to find in every cell of a dog's body, and secondly, how many of these letters you think are going to be different when you compare two dogs from different breeds. And then we'll come back and discuss the answers.

[00:09:50.59] OK, now let's see how close you got with your answers to what we actually see with the dog DNA. It turns out that in every single cell of the dog, you've got 2.4 billion letters of DNA. And when you compare two dogs from the same breed and you look at their DNA, you're going to find about one change every 16 hundred letters. And so for the most part, they're going to be exactly the same at every position, but you do see this very small number of changes. With one change every 16 hundred letters, that would mean that with a 2.4 billion base genome, you have about 1.5 million differences between any two dogs. In order to map a gene, though, we don't have to look at all 1.5 million differences, we only need to look at about 1 in 10 of those. There are about 170 thousand. And we do that using something that we call a SNP chip that I'm going to show you more about later on.

[00:10:43.01] So now we're going to actually map a gene. So the one we're going to map is this trait called white coat color in the Boxer breed. These are two very cute Boxer puppies and you can see that they're two different colors. One of them is entirely white and one of them's what we call solid, meaning that he has color almost all over his body. Some of the solid dogs are brown but the one I'm showing you here's a solid dog that's black. And what we're going to do is we're going to compare the white dogs to the solid dogs and try to find the gene that's different between them. So one other thing I want to mention about this trait before we go on that's going to be quite important later on, is that if you look at the white dogs, for the most part they're all white, but about 5 out of 100 of them are also deaf, they have trouble hearing. And this is going to be important when we actually find the gene at the end of this lesson.

[00:11:30.06] So what we're going to do is we're going to take 15 solid colored dogs and 15 white dogs and compare them. And in order to do this there's going to be five steps involved. The first thing that we need to do is get a blood sample from each of our dogs and we do that by working with veterinarians. And they do it exactly the same way that you get a blood sample from a person. If you've ever had a blood sample drawn, they usually use a needle and they often get it out of the inside of your arm. And they do exactly the same thing with dogs. They use the front leg of the dog and they take it out of the inside of their leg. And the dog goes off afterwards and is usually perfectly happy, especially if you give the dog a treat, they're very happy then.

[00:12:07.66] And so then you've got your blood sample, and what you need to do is get the DNA out of the nucleuses in all of those cells. And the way you do that is using pipets. So what I'm showing you here is a pro channel pipetter, which lets you handle 12 samples at once. And we're going to use these disposable

pipet tips to handle the DNA samples. So we've sheared the DNA, so we've extracted the DNA from the cells by basically bursting the nuclei to get it all out, and then spinning the DNA in order to get it out of the cell mixture. And now we're going to take our DNA like this in our 12 channel pipetter, pull up the amount of DNA that we want, and load it into this 96 well plate.

[00:12:47.35] And so now I've loaded in 12 samples, you can see that this 96 well plate can actually hold 96 samples. And each of those samples, once it's been loaded into the plate, can then be put onto one of our SNP arrays. And these are the SNP arrays, they're also called SNP chips, and these SNP chips can basically test 170 thousand SNPs in each of our dog DNA samples.

[00:13:11.67] Now as I mentioned at the beginning, we often do studies that don't involve 15 dogs, or 15 solid dogs and 15 white dogs. But we do have samples, we have studies we're doing with hundreds and hundreds of dogs. And in those cases we can't do it by hand, even with a 12 channel pipetter. It's going to take too long, and so instead we use robots, and we use what we call a fluid handler. And this is basically a robot that uses pipet tips just like I just showed you. And it then does basically what I showed you, but because the robot's doing it, it can do a whole lot more of them. And it takes each of the DNA samples and loads it into these 96 well plates so that they can be run on the chips.

[00:13:51.93] And so at this point we've gone through our list, and we've collected the blood sample from the dogs, we've extracted the DNA, and we've put the DNA onto the chip in order to measure the SNPs. And now we actually need to put the chip onto the scanner. Now that we've loaded the DNA onto these SNP chips, we need to run it through a scanner that'll tell us what each dog has on each of these 170 thousand SNPs.

[00:14:17.04] And so what you're seeing here is a robot arm. This is part of a robot that we use to load our chips into the scanners. Because when we're doing these studies that have hundreds and hundreds of samples in it, we don't want to do this by hand, and also because robots are much more precise and much more careful than any person could ever be. So now you're going to see the robot arm is going down, it's going to pick up a plate, and it's going to load this plate into the scanner. And what a scanner does is it's basically like every other scanner you've used to scan in images, except what this scanner can do is it can scan the very tiny little 170 thousand dots at each SNP. Here you see the three chips going into the scanner. So that's three chips, three dogs' worth of DNA, and then this is the scanner. And it's got all of these plates that it's feeding in, and it's going to give us a computer file that has our data in it.

[00:15:04.50] So now we've run our chips on the scanner and we've gotten our data out. And what I'm going to do now is give you an exercise to do that's going to give you a better idea of what it is that we're looking for in the data that we've gotten from our Solid Boxers and our White Boxers. So if you think about it, what we've done-- so this is a simplified example of what we're doing here-- I'm showing you four Solid Boxers at the top and four White Boxers at the bottom. And for the most part they look exactly the same, just like I mentioned at the beginning. Most of the places in the genome the dogs are going to look exactly the same as one another.

[00:15:34.42] But there's a few places where we see differences. And the first position right here is what we call 100% correlated, meaning that all of the solid dogs, as you can see, are T and all of the white dogs are A. So the solid dogs look completely different from the white dogs, all of them. You're also going to see positions where you see differences that are not 100% correlated, they're somewhat correlated, meaning that the solid dogs look different from the white dogs, but not completely different. Here three of the white dogs have an A but one of them has a G. Notice the solid dogs, three of them have a G but one of them has an A, so they're not completely different from the white dogs.

[00:16:10.41] So now I'm going to give you this worksheet and what you're going to want to do is fill in every empty cell on this worksheet. So in the first column, the one that's labeled Position 1, in each of the empty boxes you're going to fill in the letter to kind of create the pattern that you'd expect to see at a 100% correlated SNP. In the second column, you're going to fill in the empty boxes to create the pattern you'd expect to see in a SNP that was completely not correlated with the trait you were interested in. And in the

third column you're going to fill in the boxes to make it look like a SNP where it was partially correlated. And then we'll come back and take a look at what your answers are.

[00:16:52.79] Now that you've had a chance to fill out the worksheet, let's take a look at what the answers are on it. So if we look at the 100% correlated SNP in the first column, you want to fill in the boxes so that all of the white dogs look the same and all of the white dogs look different from the solid dogs. So you can see here that all of the white dogs have an A and now all of the solid dogs have a G. So they're completely different from one another, so this position is 100% correlated with the trait.

[00:17:18.85] The second position is 100% uncorrelated. And so you basically don't want to see any difference at all between the white dogs and the solid dogs. And so if you fill in the white dogs with G's, and the solid dogs with T's, now you can count and say, all right, two of my white dogs have a T and two of my solid dogs have a T. And then two of my white dogs have a G, and two of my solid dogs have a G. So the solid dogs and the white dogs look exactly the same, so it's 100% uncorrelated.

[00:17:45.92] And then the final column is the partially correlated. So this is where the white dogs look a little bit different from the solid dogs but not completely different. And your answers here could be slightly different from the answers that I'm giving and you would still be right. So this is a little bit more variation in the answers here. But you could fill it in like this. And so here I filled it in so that three of the white dogs have a G, and one of the white dogs has a C, and then one of the solid dogs has a G, and three of the solid dogs have a C. So there is a difference between the white dogs and the solid dogs. The solid dogs have more C's but there is a C in the white dog, so it's not 100% correlated, it's partially correlated.

[00:18:25.80] So this is basically what we need to do with our data. So if you remember, we got our blood samples from the dogs, we extracted the DNA, and then we ran it on the machine to get at every single one of those 170 thousand SNPs. Figure out what the white dogs looked like, and what the solid dogs looked like, and now we want to compare them and figure out in each SNP how correlated is that SNP with the trait of being white or solid colored.

[00:18:48.42] And what we're going to get out of this association when we do this comparison, is each of the SNPs is going to get a score. And the SNPs that are 100% correlated are going to have very high scores, the SNPs that are 100% uncorrelated will have very low scores, and then we'll have SNPs that have kind of middle range scores and those are the partially correlated SNPs. And the way that we're going to run this association, the way we're going to actually analyze the data, is using a program called PLINK.

[00:19:13.82] So now I'm just going to show you how I would run PLINK and this is actually real data from Boxer dogs. Some of them are white and some of them are solid. And let's first look at the data that we got off of the machine that looks at all of our SNPs. And this is what it looks like. This is data for one of the dogs and you can just see that it's telling us what it has all the way across its genomes. There's A's, there's G's, there's T's, there's all sorts of different things. And now we've started a second dog and this dog is called Boxer 94.

[00:19:40.34] So that's what our data looks like. Now we're going to run PLINK. And so we run PLINK, it's called a command line program, so I just run it here by typing in the command line. And it runs and it does our association, it tests all of our SNPs, and it basically scores them all for how correlated they are. And now we're going to go and look at those scores on a graph.

[00:20:01.81] And so this is the graph of the data from the Boxers, and so this is real data for real Boxer dogs. And if we look across the genome, what we're seeing is that a lot of the SNPs have very low scores and then every once in a while, you can see there's that pink chromosome that's labeled eight and there's a purple chromosome that's labeled 11 and they have kind of half-height little peaks showing up. And those are the regions of the genome that are partially correlated. But the really exciting thing about this chart when you're looking for the gene that has to do with white coat color, is we have one really, really big peak of SNPs. And so you can see on Chromosome 20-- that's the green chromosome-- we have a very tall pile

of SNPs. And so each one of those dots represents a SNP, and then each one of those SNPs has a score. And so the higher the SNP is than the bigger the score is, and the more correlated it is.

[00:20:48.79] So what this chart is telling us is that there's a region on Chromosome 20 that holds the gene that's making these dogs white or solid. But it doesn't tell us what the gene is. And that's what you are going to do now as a class. And so as a class you're going to fill in this big chart. And so you're going to break up into 15 different groups, and each group is going to fill in one row on this chart. And so you'll each get a worksheet, all of your worksheets are going to be slightly different from one another, and I'm just going to show you here what the Group 1 would fill in on the chart as an example. And so they have a worksheet here that's labeled Solid Boxer 1 and White Boxer 1. So they're going to fill out the first row of the chart, the one that's labeled Pair #1. And at the first position you can see that the Solid Boxer and the White Boxer look different from each other. So then in the first empty box on that big chart they're going to fill in what the white dog has. Because the two are different and so they're going to fill in a T because that's the White Boxer.

[00:21:48.43] At the next position the White Boxer and the Solid Boxer look the same, and so they're actually going to just leave that box empty. When you go to the next one, they're different, and so they're going to fill in what the White Boxer has, so a T. And then, same thing at the next one. The Solid Boxer's an A, the White Boxer's a C, so they'll put a C in. And then going all the way across to the end, they'll fill out the whole row. So there's some of the boxes that are empty, which is where the Solid Boxer and the White Boxer look the same. And then when there's differences, they're filling in what the White Boxer has. So now as a class I want you to go and fill in this whole chart. So each group is going to fill in one pair and then we're going to come back and see what the data is telling us about the gene for white coat color.

[00:22:39.58] Now that you've had time as a class to fill in that big chart of positions on Chromosome 20, let's take a look at what it means. So once you filled the chart in, it's going to look something like this. What you're going to see is you've got a whole lot of letters here but there's only one position, which is 100% correlated with the white coat color trait. And that's the position right here. You can see that at that position all of the white dogs have a G and that's different from the solid dogs, because you're only filling in the positions where the white dogs are different from the solid dogs. If you go one column to the left of that, nearly all the white dogs are a G there. But you can see you've got one white dog that looks just like the solid dog and the block hasn't been filled in. And if you go one column to the right, you can see that for the most part the white dogs do look very similar to one another, but the top dog, the first row, the white dog looks like the solid dog and the square hasn't been filled in. And so we just have this one position where all of the white dogs looks different from the solid dogs.

[00:23:36.71] And that's telling us that this is the region in the genome where the gene that's controlling white coat color sits. It turns that there's a free program available on the Internet that anybody can use that allows you to look in the dog genome and say 'what is in this region of the dog genome?'. And we're going to go to that now. So the genome is all of the DNA that's in a dog. And a genome browser, it's this free program on the internet. This is a website, genome.ucsc.edu, that let's us look in the genome at that point. And so it does lots of different animals, we're interested in mammals here but we're not interested in Orangutans, we're interested in dogs. We'll go to the dog genome. And then, we're going to go to -- our gene was on Chromosome 20, if you remember, that's where our peak was. And then, the positions that we're interested in.

[00:24:27.56] And let's see what's there. And so this is that position of the dog genome, it's about 160 thousand letters of the genome. And you can see that it says, when we look down here, it says, over and over again, MITF, MITF, MITF, MITF, MITF. And that's a gene called MITF. There's a lot of different versions of it in here. You can see the first word is varying here, and that just tells you what species we're looking in here. There's mouse, moose, and then Homo sapiens for human, Gallus for chicken, and Rattus for rat.

[00:24:53.25] So we've looked in this region of the genome. There's only one gene, it's called MITF. It's got something to do with the trait that we're interested in. But what does it do? So now we're just going to click on one of the genes here and we've got a nice description here that tells us what this gene is doing. And this

is the part of the description that tells us what happens when there's mutations in this gene, because we expect that there's been a change in this gene in the white dogs. It's making them white. So what does this phrase mean?

[00:25:20.78] So what the phrase says, is that mutations in this gene cause auditory and pigmentary syndromes. And those two words are very important. So pigmentary, that has to do with pigment or color, and that's obviously related to the trait we're interested in. But auditory means that it has something to do with hearing. And if you remember from when I showed you this cute picture of the puppies much earlier in the lesson, what I said was that not only are all of the white dogs white, but a very small percentage of them, about 5 out of 100, also has hearing problems. And this is really exciting because that means the region of the genome that we found, that has the gene that we think is connected to the trait we're interested in -- this one called MITF -- fits perfectly with the trait we're looking at. Not only are the dogs white, which means it has something to do with pigment, but they're occasionally deaf, which means that it's got something to do with auditory, as well.

[00:26:09.58] And so we've mapped our gene and it's called MITF. So if you think back to where we started here, we took the DNA out of the dogs and we tested 170 thousand SNPs in the DNA of each dog, compared all 170 thousand of those SNPs, found that we had a very strong correlation, very high scores at this region on Chromosome 20. And then look much closer at it to find out exactly which SNP was the very best SNP, the most correlated SNP in that region, and then found out that there was one gene that sat near it, this gene called MITF. And so we found the gene that's causing white coat color in these Boxer dogs. And this is exactly the same process that I use in all of my research where we're looking for genes that are causing something. White coat color we looked at, but we've also been looking at cancer, and we're looking at epilepsy, and we're looking at diabetes. We're trying to find all of the different genes that are causing these diseases in dogs.

[00:27:02.79] So thank you very much for joining me today for this lesson and I hope you've learned something about the genetics of dogs, and also about what makes dogs special.

[00:27:22.64] Hi, thank you very much for your interest in my lesson on mapping genes and dogs. In this lesson, your students will get a chance to search through the entire genome of the dog and find a gene that's causing a trait. In this case, white coat color. And in this lesson they're going to see footage from the lab that shows robots and machines that we use to actually do this work, as well as do the actual testing for correlation to find the position in the genome that's associated with the trait. And then you go into the worldwide web and using a genome browser to find the gene.

[00:27:59.96] It's probably helpful if your students know something about DNA before the lesson. The idea that it's made up of A's, C's, T's, and G's, and the differences in DNA in people sometimes cause diseases. As far as supplies, you probably need nothing much more than just paper, writing utensils, and the ability to print out or provide the handouts that go along with this lesson. And a board at the front of the room would also be helpful.

[00:28:27.02] So it's divided into seven parts. The first part is just a very brief introduction where I introduced the idea of dog breeds and then give the students a short lesson talking about which of these dogs is bred for which of these behaviors. It's just an introduction for the idea of it. And then in the second part I start talking about where dog breeds come from, and it just ends with a handout that's "how do you make a dog breed" example. And then in the third part I discuss why the process of creating these dog breeds also leads to higher rates of diseases in these dog breeds. And then finish that off with asking them to brainstorm about which diseases they think that dogs are going to be getting.

[00:29:10.06] And so then after they've brainstormed as a class, the fourth part starts out with the actual list of the dogs' diseases that are most common in dogs. And it turns out that these diseases are exactly the same diseases that people are getting, things like cancer and diabetes, which is why dogs are a really cool species to be looking in for disease genes. And then part four kind of ends just talking about why we're using dog breeds at all, if they're the same as humans. Dog breeds, it turns out, it's really easy to map genes

in them, because dogs in a breed are so similar to each other. So that ties back to the discussion at the beginning of the lesson about how breeds were created.

[00:29:45.85] And then I just talk very briefly about what the genome is. I avoid using the term genome, I just talk about the DNA that's in a cell and how much of it there is. And then end off that part asking the students to brainstorm about how much DNA do they think there is in the cell, how many letters, and then how many differences they think there's going to be between dogs.

[00:30:04.05] And so then in part five, this is where we kind of get into exactly how we're going to actually map this trait. And so I give the trait that we're going to map. I describe it, it's a white coat color trait. It shows some cute puppies, what's different between the solid colored puppies and the white puppies. And then I explain that in order to find the genes that we're interested in, we look at 170 thousand different SNPs, or differences, positions in the genome that are different between dogs. We look at 170 thousand of these positions and find the positions most correlated to the gene, or, sorry, the position that's most correlated to the trait we're interested in. And then that part five ends off with a worksheet that just introduces the concept of correlations and gets the students to-- basically it's a worksheet that says "what does it look like if you're 100% correlated?", "what does it look like if you're not correlated?" And just to really make sure that they understand what a correlation is and what it is that we're looking for .

[00:31:02.84] So then in part six we go over this worksheet on correlation, and then we show how we're using this with the data that we generated for these 170 thousand SNPs. So I should mention that in part five, that's also the part where we go over the lab work that's involved, and have some footage of the robots that are working in the lab and things like that. Some shots of pipetting that we use to extract the DNA from the dog's blood samples. But anyway, so in part six, we're looking at actually mapping the locus. So we look at the data, we run it through the program that we use to find the association, we look at a graph that shows the correlation, and then finally we end off with a worksheet that-- basically at this point we've gotten it close to the gene, we haven't pinpointed the gene of interest, and this worksheet is what they're going to do to actually pinpoint the position in the genome that's where this gene is.

[00:31:50.23] And so this isn't actually a worksheet, this is a group exercise for the entire class, where there's a big chart that's got 15 rows in it, and then the students break up into 15 groups. And each of those groups fills that one row in the chart. So you could draw the chart on the board, you could project the chart on the board if you had a projector, or on a whiteboard. It just gets each group of students, they each have a handout, and each of their worksheets is slightly different from one another so you do have to fill out all 15 rows of that table in order to get the exercise to work. Because then, after they filled in the chart -- and this is part seven -- you'll find out that when the chart's completely filled in there's only one position in the chart that's 100% correlated with a trait we're interested in. And at that point we know where in the genome the gene we're interested in is.

[00:32:32.53] I show going to the genome browser, looking up the website that can tell us what's in the genome in that position. We look at that position in the genome and there's only one gene called MITF. And that's our white coat color gene and that's where the exercise ends.

[00:32:45.76] And so basically, it kind of starts with why dog breeds are cool, takes you all the way through the process of mapping a gene, which is exactly the same process that they use in human disease studies that you may read about in the news where they're studying things like cancer and other diseases in humans, but we're just doing it in dogs. And then takes you all the way to finding an actual gene, and I should also mention here that the example we're dealing with all the way through here, white coat color, is a real example and we are looking at real data. So this isn't anything that's been made up, this is the study that we published a few years ago, so it's a very concrete example of what we do. So anyway, thank you very much for your interest in my lesson and I hope your students enjoy it. I think it's a lot of fun. Dogs are always fun to talk about and it's a great example of some really cutting edge science that's going on around the world today. So thank you very much.

[00:33:40.48] [MUSIC PLAYING]

