An Introduction to Evolutionary Biology

Prof. Sutirth Dey

Biology Department, Population Biology Lab

Indian Institute of Science Education and Research (IISER) Pune

Week 3 Lecture 13

From family to the population

So, in our last discussion, we derived a very simple relationship: p = P + Q/2. And we said that we were going to require that to extend Mendel's laws to the population level. So, let us do that in this discussion. So, let us assume that we have a diploid, sexually reproducing population. And we are dealing with a scenario that is a one-locus two-allele case. So, a trait that is determined by a single gene at that particular locus has two alleles called A1 and A2.

So, let us assume that the frequency of the allele A1 in the parents is p. Now, from this frequency of A1 in parents, We need to go to the frequency of A1 in all the possible gametes that will be formed. And if certain conditions are met, if segregation is normal, equal fertility, no mutation, no migration, then We claim that the frequency of A1 remains p. Now, why is that so? Now, think about all the ways in which the frequency can differ between the parents and at the gametic stage.

If, for example, when gametogenesis is happening, the segregation of the alleles is not equal, Not the Mendelian segregation that we are aware of; in that case, this frequency will be incorrect. Similarly, suppose one of the genotypes is producing a greater number of gametes compared to the other genotypes. Say A1A1, let us say, is producing more gametes than A2A2. In which case, obviously, A1 will be more represented in the

population, and the frequency of, you know, A1, p will go up. So, we need equal fertility to ensure that P does not change.

Similarly, suppose there is a mutation happening; again, that will lead to a change in p. Or suppose there is migration happening; you know some individuals are leaving the population. If some other individuals are coming in, again that will lead to a change in the frequency of A1. And that is why, in the absence of all these changes, The frequency of A1 in the gametes is going to remain the same as it was in the parents, which is p. Now, remember that in any organism, the total number of gametes produced is much larger than the actual number of gametes, which end up forming zygotes and leading to the formation of babies. Therefore, from this large number of gametes, we have to select those gametes that are forming the zygote. And we are claiming that if the population size is large and if all the fertilizing capacities of all the gametes are similar, Then the frequency of A1 in gametes forming zygotes will still remain p. Now, why are we making this claim? Remember, we are going from the frequency of A1 in all gametes to the frequency of A1 in a subset of the gametes, right? So, basically, there is sampling happening. Now, whenever a sampling happens, the larger the sample and you know the more uniformly and unbiasedly the sample has been taken, The more the chances are that the frequency in the sample is going to be the same as the frequency in the population. Now, let us look at the first part: the large sample size thing. How large should a sample be? Obviously, it should ideally be infinitely large. Now, in order for a sample to be infinitely large, The population from which the sample has been taken also has to be infinitely large. So, that is why, technically speaking, we say that this is an infinitely large population.

Now, similarly, I said that the sampling had to be unbiased. In other words, all the individuals in the population should have an equal chance of being represented in the sample. Now, if that is the case, then in a biological system, what does it imply? That implies that all the individuals or gametes have an equal chance. of being ending up being the gametes that are forming the zygotes. And that is ensured only when the fertilizing capacity of all the gametes is similar.

So, for example, suppose the gametes bearing, say, A1 have a greater fertilizing capacity for whatever reason than the gametes bearing A2, What will happen? Obviously, the frequency of A1, which is p, is going to increase. Similarly, if you know that A1 has a lower fertilizing capacity, p will decrease in the gametes forming zygotes. And as long as the population size is very large, infinitely large, and the fertilizing capacity of all the gametes is equal, We do not expect this particular frequency to change. So, it remains at p. Now, from the frequency of A1 in gametes forming zygotes, we need to go to the genotypic frequency among all the zygotes.

How do we get there? Now, in order to go there, we will have to make two major assumptions. What are the major assumptions? A. Mating is random, and B. the frequency of males and females is equal. Now, why are we making these assumptions? So, if you remember, we said that.

The process of forming the zygotes in the next generation is essentially a sampling from two gamete pools. The male gamete pool and the female gamete pool. In other words, all you can think of the process of formation of the zygotes as putting your hand in the male bag. Bringing out any one gamete, putting your hand in the female bag, bringing out any gamete, and then simply putting them together. Now, suppose that these two draws are independent of each other; in other words, If, let us say, you have put your hand in the male bank and you have got an A1 gamete, Does that affect what gamete you are going to get from the female bank? If it does not, then in the language of probability, we say that the two events are independent.

Now, if the two probabilities are independent, then you have a very nice result from probability theory. which states that if two probabilities are independent, Then the probability that both of them are happening together is essentially the product of the two. So, what we can do is essentially make what is known as the Punnett square. Let me make it here so that I can make it a little larger, and you can see it better. So, we create what is known as the Punnett square.

So, let us assume that this side is male. So, you have A1 and A2. Here are my two types of gametes. Now, the males are producing. This is being produced at frequency p; this is being produced at frequency (1-p) or q.

And similarly, let us assume that this is the female side; this is my A1, and this is my A2. Now, if we assume that the frequency of the gametes in the males and females is the same, That means that this is also going to be p, and if this is p, this is (1-p), which means this is also going to be q. So, this is where the equal-frequency assumption comes in. Now, what is the probability of an offspring being A1, A1? So, the probability is putting your hand in the male bag and pulling out an A1 gamete, the probability for which is p. and then putting your hand again in the female bag and pulling out a A1 gamete, the probability for which is again p, this p.

So, the probability that you will form an A1A1 zygote is that you know both these events are happening together. which essentially means that the probability of formation of A1A1 is (p * p), which is p^2. Similarly, the probability of forming an A2A2 gamete and an A2A2 zygote is (q*q) = q^2. Now, what is the probability of forming a heterozygote? Now, you can form a heterozygote either by pulling an A2 gamete from a male and an A1 gamete from a female, The probability for which is going to be (q*p), or you can do it the other way around, in which case the probability is going to be (p*q). Now, assuming that it does not matter whether you are getting your A1 from your mother or your father, These two, you can consider them pq = qp.

In other words, the three genotypes that you are going to get, These are going to be in the ratios A1A1 : A1A2 : A2A2; these will be in the ratios p^2, 2pq, q^2. That is exactly what I am telling you here: the genotypic frequency among the progenies will have to be this. Now, what happens if the two conditions that we are putting forth are not met? If the mating is non-random, then this nice multiplication that we are doing is not going to happen. Because that random mating is what is ensuring the probabilities of drawing the gamuts from the male. The male and female pools are independent of each other, and

because they are independent, you are able to multiply them this way.

If mating is non-random, then this simple multiplication will not work. And what happens if the frequency in males and females is not equal? Then, of course, if you have p and q over here, you will have some other values in the female part, right? In which case, again, the overall logic will remain correct, but it will not be (p*p); it will be p into something else. It will not be (p*q); it will be p into something else, and so on and so forth. So, as long as you have random meetings and as long as you have equal frequency in both males and females, You are going to get your ratios; you know, your genotypic frequencies in these numbers, or rather these values p^2, 2pq, and q^2. Now, your zygotes have been formed, but not all of the zygotes are going to be able to survive to become adults.

So, the genotypic frequency among the progeny need not necessarily be the same as the genotypic frequency among the zygotes. However, if all the zygotes suffer equal viability, in other words, all of them have similar survivorship. which you know equally means a similar rate of mortality; then, even if some fraction is dying off, As long as the fraction is the same, these relative ratios p^2, 2pq, and q^2 are actually going to remain the same. The relative frequencies—sorry, these relative frequencies—are still going to remain the same. And now remember, we started with the frequency of A1 in parents; we have to go to the frequency of A1 in the progeny, right? Now, how do we go from the genotypic frequency in the progeny to the frequency of A1 in the progeny? So, if you remember, I told you that small p, which is the frequency of A1 in the progeny, is going to be P + Q/2.

Now, here P is the frequency of the A1A1 genotype, which means it is equal to p^2. And Q/2, so Q is essentially 2pq. So this is going to be equal to [(1/2)*2pq], which basically means it is going to be $p^2 + pq$. Now, you can call this p dash or p' (prime) to show that this is in the progeny generation. So, this p' is now equal to p, you can take this common into (p + q), right? $p' = p(p + q)$ But q we already know is 1 - p.

So, (p + q) = 1, right? Therefore, this entire thing becomes equal to p, which is what I am

showing you over here. As long as there is equal viability among all the progenies, this is going to remain as p. In other words, if you are starting with the frequency of A1 as p, then assuming all these conditions are met in the progeny. The frequency of A1 is still going to remain at p; it will remain unchanged. And furthermore, there is a certain, you know, relationship between the frequency of A1 and the gene frequency of A1.

The genotypic frequency and that relationship are given by this number; or rather, this is the relationship. which is p^2, $2pq$, q^2, which essentially means these are the terms of the square of $(p+q)$. This leads us to the famous Hardy-Weinberg law. Which was propounded by these two gentlemen, Godfrey Harold Hardy from England and Wilhelm Weinberg from Germany. Now this is the same Hardy who ended up discovering and mentoring Ramanujan.

So, obviously, he has a very nice connection to India, but I mean, although we Indians tend to think of him as Ramanujan's mentor, He also has a very important position in the, you know, annals of mathematics in the sense that he is a very famous mathematician. Now, whatever we have discussed so far, we can summarize it in the form of a statement. In a large random mating population with no selection, mutation, or migration, The gene frequencies and the genotypic frequencies are constant from generation to generation; they do not change. And there is a simple relationship between the gene frequencies and the genotypic frequencies. Which, as I showed you in the 1 locus 2 allele case, is given as this.

In other words, the ratios of the various genotypes are the terms of this particular expansion. Now there are quite a few comments that I need to make in the context of what we just saw. The first comment is that I mean this is something that many of you are probably thinking. Isn't this a bit of cheating because what we are doing is seeing that at every single stage? What are the ways in which a difference can occur in terms of the frequencies? We are explicitly going out of our way to say that all the conditions are such that there will be no difference. And therefore, after all the conditions of no difference are put in place, what do you say you see no difference? In other words everything that could

have made a difference are nullified And then you get a great result saying, "Hey, look, allele frequency is not changing; unit frequency is not changing.

I mean, why is that a thing? That is a very valid question, and the answer to that question lies in the way in which this entire relationship was discovered. So, it turns out that if you remember, you know when Darwin came up with his theory of evolution through natural selection. The requirement of inheritance was very important because, without inheritance, natural selection could not have worked. And then, of course, Darwin did not have a way, you know, to explain inheritance, and that is where Mendel's theory comes in. However, in Mendel's theory, you have this idea that in the F2 generation of the monohybrid cross, you get a 3:1 phenotypic ratio.

In other words, you know 3 parts will be tall, three-fourths of the population will be tall, and only one-fourth of the population will be short. Now there was a particular meeting that happened, I forget where, somewhere in England, in which W. C. Punnett, A very famous geneticist, the same person on whom that Punnett square that I just now made has been named. W. C. Punnett was giving a lecture, and in that lecture, One of the people who was attending was a very famous statistician called Udny Yule. Now, Udny Yule makes a comment during that lecture. He essentially says, or at least that is what Panett thought, that he is saying, "Look." If you are going to have a 3:1 ratio of tall:short or of dominant:recessive phenotype, then over time, The dominant phenotype is going to take over the population; the recessive phenotype is essentially going to go extinct. In other words, over time only the dominant gene is going to be present in the population; the recessive gene is just going to disappear.

And if that happens, then remember that without variation, evolution cannot proceed; without variation, natural selection cannot proceed. So, therefore, he claimed that Mendelian theory, when extended to the population, then the outcome of that is going to be that all the genetic variation on its own just disappears. Now scholars are of two opinions about whether Udny Yule actually said this or meant something else. There are many people who think that you know Yule meant something else. I mean, he was a very

famous statistician; he was a very sharp person.

However, what is not doubted is that this is what Punnett understood. Now, Punnett was not very happy with the whole situation. He realized that there was some kind of fallacy happening somewhere, but he did not really understand what the hell was going on. Now, Punnett used to play cricket with Hardy. So, he comes back and has a chat with Hardy, where he explains this entire thing to him.

Udny Yule's objection states, "Hey, I do not know what is going on; can you help me out here?" So, Hardy, being the very sharp and famous mathematician that he was, immediately understood what the problem was. And therefore, he immediately sketches out the proof of what we just described on the back of the envelope. It does not really matter whether a trait is, you know, recessive or dominant. As long as certain conditions are met, those are the conditions that we discussed. As long as those conditions are met, the gene frequencies and the genotypic frequencies will not change over generations.

In other words, the raw materials for evolution will always be there. Now, this was a very significant result. This suggested that the variation by itself will not disappear. Only when evolutionary forces act, only then will the frequencies change, only then will evolution happen. In the absence of those, you know, like the law of inertia, The population will continue to have the same amount of variation that it used to have.

Now, Punnett himself was a very sharp individual. He immediately understood the implications of this entire situation. However, there was a problem: what was the problem? The problem was that Hardy considered himself to be a "pure mathematician." In other words, he was of the opinion that mathematics should be done only for its own sake. Mathematics should not be done just to know that it leads to a benefit or to prove and make some applied contribution. Therefore, Hardy was extremely reluctant to get this thing published because he thought this was an absolutely trivial result.

And more importantly, it will make life useful for somebody, the biologist, whom he

obviously thought was beneath his stature. But Punnett ended up, you know, pestering him, and finally, At some point, Hardy relents, writes his paper, and ends up publishing it in the journal. So, this is a very, very famous paper—just one page—and it starts with this absolutely brilliant thing. I am reluctant to intrude in a discussion concerning matters about which I have no expert knowledge. I should have expected the very simple point that I wish to make to be familiar to biologists.

So, you can see the innate bias or innate superiority complex of being a mathematician showing. Then he goes on to say; however, some remarks of Mr. Udny Yule, to which Mr. R. C. Pannett has called my attention, suggest that It may still be worth making, and then he goes on to sketch that proof. Now, this is the sole contribution that Hardy makes to the field of biology, and very interestingly, You know, of course, Hardy, if you talk to the mathematicians, they will tell you what Hardy has done. However, outside the world of mathematics, Hardy is actually way more famous for making this certain contribution in the context of the Hardy-Weinberg law. which is a big irony given how Hardy would have completely hated the fact that you know His work made some important, application-oriented contributions. Anyway, Hardy writes his paper in English, as you can see; therefore, This becomes reasonably well-known in the English-speaking world.

However, around the same time, in 1909, actually, Wilhelm Weinberg completely independently discovers the same principle. In fact, he ends up doing a lot more work related to this than Hardy ever did. I mean, Hardy did nothing; he just published this paper and stopped. Weinberg does a lot of work, but unfortunately, he was publishing in German. That is why his contributions were not known to the English-speaking world for quite some time.

And at some point, when it became clear that Weinberg had figured it out on his own and had done a lot more, At that point, it was proposed that the names of Hardy and Weinberg should be put together. Both of them should be given credit for this discovery, and that is how we got the Hardy-Weinberg equilibrium. Now, this statement is the complete statement of Hardy-Weinberg equilibrium. However, it has been my observation over the

last several years that Students somehow do not think of this statement as the Hardy-Weinberg equilibrium.

So, whenever I ask people in the interviews, etc., what the Hardy-Weinberg equilibrium is, Then they simply say Hardy-Weinberg equilibrium is $(p + q) = 1$ and $(p^2 + 2pq + q^2) = 1$. Now, the problem with these two statements is that they are correct. In a 1-locus 2-allele scenario, the genotype frequencies and the allele frequencies will sum up to 1; there is no doubt about that. But the point is you do not need the names of two great mathematicians to be associated with a result if that were only trivially true. And that is why I tend to sometimes ask students, you know, if they are saying that $(p+q)=1$; $(p^2+2pq+q^2) = 1$.

Why do you think the names of two great people are associated with such a trivial result? And you know what the answer is that I get? So, one of the answers I received was, "Yes, sir, I fully agree it is a trivial result." But you see, even a trivial result needs to be thought of the first time, and these were the two people who thought about it the first time. which is why you know their names are associated with it. This, incidentally, as you can see, is not correct; this is not the statement of Hardy Weinberg; this is what Hardy Weinberg stands for. The other very important thing to note is that if this is all there is to it, this statement, then, Why is it that I am showing you this entire, you know, derivation? Why is it that I am pointing out where each and every one of those assumptions is impinging? The reason for that is that, by itself, the Hardy-Weinberg law is absolutely uninteresting.

As I said, the Hardy-Weinberg principle is simply stating that If all the ways in which allele frequencies can change are blocked, then allele frequencies will not change, right? So, in that sense, the law itself is not interesting. What is interesting is the fact that in real life it is very rare that A population is going to, you know, meet all these assumptions that I have written over here. So, if that is the case, the entire subject of population genetics proceeds by figuring out in what way All these assumptions that I have written over here, in what way are these assumptions being violated? So, in that sense, the Hardy-Weinberg law is a null law; it is equivalent to Newton's first law. What does Newton's first law say?

It states that A body is going to remain in its state of inertia, either moving or static, until and unless an external force acts upon it.

So, Hardy-Weinberg also says the same thing that. A population is going to maintain its variation until and unless an external evolutionary force acts on it. And therefore, the whole idea of population genetics, the whole goal of population genetics, is to figure out Which one of these forces is acting, and in what way are they acting? What are the relative contributions of the various forces, and so on and so forth? So, that is why this is one of the few laws in biology. Remember, there are very few laws in biology. This is one of maybe four or five that exist.

So, there are a few points to note about the Hardy-Weinberg equilibrium. The first point, which is a part of the statement itself, is that the gene and the genotypic frequencies They do not really change from generation to generation. The second point is that. The relationship between gene frequency and genotypic frequency is reached in a single generation in the one-locus case. This is very important. So, what do I mean by this? So, suppose you have the population which is in p; you know the allele frequency of one of the alleles is p which means that its frequencies, genotypic frequencies, are p^2, 2pq, and q^2. Now, suppose for whatever reason this becomes the allele frequency of the first allele, p1. In which case, you expect the genotypic frequencies to be $p1^2$, 2p1q1, and $q1^2$, right? So, the question that we are asking is how long it will take for the genotypic frequencies to go from p^2, 2pq, and q^2 to p_1^2, $2p_1q_1$, and q_1^2, etcetera, etcetera. Now, it turns out that, as I said, if it is a single generation case—sorry, single locus case—then it will happen in one generation. However, within the single locus case, if the frequencies of the alleles are different in males and females, Then it will end up taking more than one generation.

And when we go to the two-locus case, Then it typically takes way more than one generation, depending on what the recombination frequency is between the loci. and depending on something known as the coefficient of linkage disequilibrium. Of course, when you are talking about two loci or more than two loci, then things like epistasis, etc., also come into the picture. Which, of course, can you know make this genotypic

frequency relationship never be attained.

But that is a much more complicated case, something that we are not going to consider as part of this particular course. So, the third point is related to progeny genotypic frequencies. They are dependent only on the parental gene frequency and not on the genotype frequency. Now, what do I mean by that? What I mean is that our small p is equal to capital P plus half q, right? So, in principle, you can get the same value of small p by altering the capital P's and the half q's. In other words, different parental genotypic frequencies can lead to the same parental gene frequency.

And what this statement insists is that in the offspring generation, The relationship between the gene frequency and the genotypic frequency depends only on the gene frequency of the parents. Not on the genotypic frequency of the parents. What do I mean by this? So let me give you an example. Suppose I have a population in which my A1A1, A1A2, and A2A2 are in these frequencies: 0.3, 0.6, and 0.1. And let us assume that in another population, the frequencies are 0.2, 0.8, and 0; we do not have any A2A2 at all. Now, what are my gene frequencies in these two cases?

In the first case, this is [0.3 + (1/2*0.6)] ||| (1/2*0.6) is 0.3 ||| which means 0.6.

In the second case, this is [0.2 + (1/2*0.8)] ||| (1/2*0.8) is 0.4 |||

which means my gene frequency is again 0.6. So, in both these cases, the gene frequencies are the same and therefore, When it comes to the genotypic frequency in the progeny generation in the next generation, in both cases, It is going to be 0.36:0.48:0.16, which is simply $(0.6)^2$, and so on and so forth. So, if you want, you can actually simulate it for yourself and convince yourself that that is how it works. Excellent. So, now there are two further points about the Hardy-Weinberg equilibrium. But before we get to those, I need to introduce you to this diagram. So, what am I doing over here? So, here we are representing the alleles as A for A1 and a for A2.

And here, the frequency of A is p, which is going from left to right. So, as you go from left to right, p is increasing. Which obviously means that the frequency of the other allele,

a, is increasing from right to left. These three lines depict the genotypic frequency of A1A1, A1A2, and A2A2. So, this one is for, I am sorry, AA; this one is for aa, and this one is for the Aa, the heterozygote.

Now, let us first concentrate on the homozygote. As you can see, when the allele frequency p is very low, the homozygote frequency is very low. But as p increases, this, you know, increases quite fast, which is obvious because this is going by the square. And then, when this frequency of the allele becomes 1, the frequency of the homozygote also becomes equal to 1; no issues. The opposite happens for the other allele; when the frequency of that allele is very high, you know, when it is at 1, then The frequency of the homozygote is 1, and as the frequency of the a allele keeps decreasing, then This frequency of the homozygote also keeps decreasing, and When this frequency becomes 0, it also becomes 0; no issues.

The interesting thing is that for both homozygotes, the maximum occurs when p=1 or when q=1. This point for q=1 and this point for p=1. But when it comes to the heterozygote, the Aa genotype, Then you can see that the maximum value of that is occurring somewhere in the middle. And for the 1 locus 2 allele scenario, this value will become maximum when p=q=0.5, and what will that maximum value be? Simple: 2 * 0.5 * 0.5. So, 0.5 * 0.5 = 0.25, so this is going to be equal to 0.5, and that is what you can see over here. So, this is 0.5. So, I am not demonstrating this to you, but if you know very simple calculus, You should be able to show, using the first derivative and the second derivative, that the maximum is indeed this point.

The value of the maxima is 0.5, and this maxima occurs when p = q = 0.5. So, that is what I am telling you here. The next point is a slightly more subtle one, and that point is related to the relative frequency of homozygotes and heterozygotes. Now, look at some points over here.

This is where p = 0.1, and obviously p^2 is very small; it is $p^2 = 0.01$. However, even at this point, you can see that this value of the corresponding heterozygotes is reasonably

high; it is close to 0.2. So, this tells us that when an allele in the population is in Hardy-Weinberg equilibrium, And when that allele is rare, then most of the copies of that allele are actually going to be inside the heterozygote. Now, why is this a worthy point to make? It is worthy because of how human or not all, but most human diseases behave. So, if you look at most diseases, not only in humans but also in many animals, actually, Most of those diseases, genetic diseases, are caused by recessive alleles.

Now, obviously, if a disease is caused by a recessive allele and that disease is leading to, you know, Some problems in terms of survivorship or in terms of how many babies are being produced. then you expect natural selection to take away you know those alleles from the population. In other words, you expect natural selection to reduce the frequency of those alleles. Now, obviously, natural selection has been acting for many, many years; therefore, you expect that over time natural selection will completely be able to weed out those alleles from the population.

Yet many of those diseases are still very much present. How does that happen? Now, there are multiple ways in which it can happen, but one of the ways that we have to remember comes from this diagram. which is that natural selection can push the frequency of that allele to a low value. But when it does so, most of the copies of that recessive allele end up in the heterozygotes. And when they end up in the heterozygotes, they are not expressed because they are recessive, right? And at that point, natural selection cannot do anything against it. And that is why, for natural selection to weed out a recessive allele through selection alone, It becomes a very, very difficult proposition once the allele has gone beyond a certain threshold.

And this is one of the reasons why many, many recessive genes or recessive alleles are still present in the population. So, that is the point I am making here. So, before we wrap up, a quick note on why the Hardy-Weinberg equilibrium is such an important concept in biology. So, the first point is that this is one of the few laws in biology. And as I said, by itself it is hardly of any interest, but the entire subject proceeds by looking at the departures from the law.

In that sense, it has the status of a null law. It is like the first law of mechanics: Newton's first law of mechanics. which states that a body will continue in its state of inertia until acted upon by an external force. Again, there is no body on which some force or another is not acting, and therefore, the law by itself makes no major prediction. But the entire field of mechanics actually proceeds by asking how the law is being violated. And in what ways are forces acting on the body such that the state of inertia is being violated or changed? Exactly the same thing happens in the context of this particular law.

You looked at the assumptions correctly; there are so many of them. And there is hardly any population that is going to obey any, or you know, all of those assumptions. Somewhere or other, the violations will happen, and when the violations happen, there will be a departure. From Hardy Weinberg, the entire subject of population genetics is trying to figure out why that departure is happening. Which of the assumptions is being violated and to what degree? And, of course, in its purest form, this law is suggesting that In the absence of any evolutionary force, allelic diversity is going to be maintained.

This itself is a very big thing because it suggests that the raw material for selection to work. For evolution to work, in fact, all else being equal will be maintained on its own. It is not that allelic diversity is going to go away, and therefore, evolution will stop. And in that sense, this is a very strong support for Darwin's theory of evolution. And although we are not showing it to you here, But this law actually expands very easily to both multiple alleles and multiple locus cases, as well as multiple loci cases.

When it comes to multiple alleles, it is actually very, very simple. So, for two alleles, the relationships, you know, the genotypic relationships are terms of this expansion $(a+b)^2$. If you have three alleles, it becomes the term of the expansion $(a+b+c)^2$.

If you have n alleles, it simply becomes $(a + b + c + ... + n)^2$. So, the terms of this expansion become the various terms of the genotypic ratio. In the context of multiple loci, it is relatively easy to take it over there. Of course, it becomes a bit more

complicated; the math becomes a bit more complicated, but it is still tractable. And then finally, from a very practical point of view, This helps us know what the frequency of the carriers of the disease is in a population.

So, who are the carriers? Carriers are the heterozygotes. Now, if you are a public health person, you need to figure out what the rate is at which, Rare diseases, relatively rare diseases, are going to surface in the population. And I mean, obviously, this is needed so that you can take appropriate measures for treating them in the hospitals and so on. And in order to understand that rate, you need to understand what the frequency of the heterozygotes is. And as we are going to show you in the next discussion, You can do that by assuming Hardy-Weinberg and then doing some relatively simple calculations. So, that kind of simple calculation and a few others is what we are going to do in the next discussion. Bye.