Evolutionary Dynamics
Supreet Saini
Chemical Engineering
Indian Institute of Technology Bombay
Week 04
Lecture 20

Thank you. Hi, everyone. Welcome to the next lecture of the course. Today, we'll discuss a concept called the distribution of fitness effects. To understand that, let's revisit the fact that when a mutation occurs,

So, let us imagine a DNA sequence where there was a nucleotide A at a position, and a mutation occurred, changing this A to a G. This is the mutational change we are discussing. This leads to, let us say, the organism that had this particular DNA sequence having a fitness of fo. After the mutation, the progeny had a fitness f'. This mutation could have one of three different effects on the progeny's fitness. The first effect could be that the mutation significantly increased fitness.

In this case, f' is greater than fo, and we say the mutation is often represented in this format, indicating that the original nucleotide was A, and the new nucleotide is G. This standard representation shows that the mutation A to G was beneficial. Alternatively, the fitness of the mutant progeny could have decreased compared to the ancestor. In this case, f' is less than fo, and we say the same mutation was deleterious. So, beneficial mutation, deleterious mutation. When we say fitness increased or decreased, we are referring to the growth rate.

As we saw in the last video, that growth rate is being taken as a proxy for fitness. So growth rate increased, growth rate decreased, or in the last scenario, what could also happen is that fitness remains the same. In which case, f' is simply equal to F naught, no change in the growth rate of the organism took place, and then we say that this mutation was a neutral mutation. So according to this classification, mutations can belong to either one of these three groups.

So the first question that we want to ask is, of all the mutations that are happening, Of all mutations happening, what percentage of them are beneficial, what percentage of them are deleterious, and what percentage of them are neutral? So somebody who thinks that, of all mutations happening, most of them will be beneficial in nature, what you are saying is that a very high percentage of mutations belong in this category, and this is relatively smaller numbers. Alternatively, somebody else might think that, of all the mutations that are happening, mutations are occurring randomly in the genome itself.

And hence, more often than not, their effect on fitness is going to be that they will make an enzyme or a protein, which has been working well, bad. And hence, as a result, something that was working well will no longer be working well. And hence, the effect of the mutation would be a deleterious effect. And hence, a large fraction of mutations would be in this category. So it sort of depends on our intuition of what we think mutations do and how they operate.

And while we don't know these exact numbers, people are working to decipher them. While we don't know the exact numbers associated with any particular organism or context, it is largely believed that beneficial mutations are fewer. And you have more deleterious mutations and more neutral mutations. And this should make sense because mutations occur randomly. So, if we have a gene sequence like this—let's say it starts with ATG and ends at TGA.

This is going to get transcribed and translated, and eventually, we will get an amino acid chain. That amino acid chain will fold into a protein, and the protein will do its job inside the cell. However, if a random mutation such as this occurs, this amino acid will change. And more often than not, the effect of a random change will be that it makes a machine that was working smoothly worse off. If we think of an enzyme—a protein that catalyzes reactions—as simply a molecular machine, the effect of introducing a random mutation into this finely working machine will be that it makes the machine worse off than before.

And hence, that's going to lead to a greater contribution toward deleterious mutations. So, again, we don't know these numbers, but these are the qualitative measures we have. And we will look at some experimental data as we move forward in the course. What we should always keep in mind in all our discussions of fitness data—because that's what we are moving toward now—is whether a mutation occurs, beneficial or not, whether it is beneficial, deleterious, or neutral—depends on the context.

And the context that we are interested in here is, if I get this off the screen, the context is in which genome the mutation is occurring. And in which environment are we discussing the effect of the mutation? Let me discuss these two aspects. Let's imagine a DNA strand, and in this organism, there is a gene X and a gene Y. Now, a mutation takes place, let us say in gene Y, and this mutation is beneficial in nature.

But this is beneficial in nature only because this individual already has gene X present in the genome. If gene X wasn't present, then this mutation would no longer be beneficial. Hence, this tells us that the effect of this mutation in Y is dependent on the genomic context. Because it is possible that if the same mutation had happened in an organism whose genome was like this—

where gene X was lost by this species and only gene Y remained—the same mutation would occur here. In this case, this mutation is deleterious.

So, the effect of the same mutation is dependent on the genomic context. When gene X is present, the same mutation is beneficial. When gene X is absent, the same mutation is deleterious. That is the first thing we want to say—we have to keep this in mind in every study we are looking at. The second thing we are saying is that suppose I have a

A gene, and in this gene, a mutation takes place. Now, the effect of this mutation could be beneficial if I am testing fitness in an environment with lactose as the carbon source. And let us give this gene a name. Let us say this is a protein called LacY, whose job is to bring lactose in. So, this is the LacY protein, and it brings lactose inside the cell.

And via the introduction of this particular mutation, LacY is made a better protein in this gene. So, An individual which is carrying this mutation is able to bring in lactose much better as compared to the ancestor which is not carrying this mutation. So, in that context, if cells are growing in an environment with lactose, this mutation is a good mutation because it allows you to bring in lactose much better as compared to the previous version your ancestors were carrying. However, if I grow this same individual in lactose,

In an environment where glucose is the carbon source, in that case, because this is a transporter which is dedicated to the transport of lactose, In a glucose environment, this mutation is completely irrelevant because LacY as a gene is not being expressed into proteins. Even if it was expressed, it is used to bring in lactose, and it is useless in the context of an organism's physiology when it is being grown in a glucose environment. In this case, this mutation is a completely neutral mutation. Because it does not give you any benefit as compared to the other.

This is a gene which is not relevant in the environment that I am talking about. The lesson for us from this discussion is that effect of a mutation is environment dependent. And that these two are lessons that as we move towards these precise studies of evolution experiments and we develop the theoretical framework in which we are going to look at them, these two things have to be kept at the back of our minds at all times. That effect of any mutation is contingent on the genomic background, which was what we discussed in this slide. But also that effect of any mutation is context dependent in the sense of which environment was the effect measured in.

The same mutation could have polar opposite effects depending on the environment in which we grew the cell. To take another example of this environment dependence, let us think of E. coli two variants. So, let us imagine that we have two individuals of E. coli and let us draw their genomes.

Again, bacterial genomes are circular. Everything is identical in these two individuals except for the fact that the individual on my right is carrying this one mutation.

So we saw that bacterial genomes are of the order of length 5 into 10 power 6. Except for this one nucleotide, they are identical everywhere else. What this one mutation does is this mutation confers resistance to an antibiotic. which seems like a very useful property to have. So, let us call this individual A, let us call this individual B. If I were to grow them together in an environment with antibiotic, grow in environment with antibiotic,

Then who do you think is going to flourish and who do you think is going to be weeded out by natural selection? Obviously, individual B, because it's resistant to antibiotics, is fitter in this particular environment as compared to individual A. So B is fitter than A. What that means, from the context of this mutation's effect on fitness, is that this mutation makes the individual fitter and hence this is a beneficial mutation in this environmental context. But now let us look at another environmental context where these two are grown, but in the environment there is no antibiotic. Often what happens is that mutations that make you resistant to antibiotics, such as this, allow you to counter the effect of the antibiotic, but when the antibiotic is not there, your growth rate is slowed down.

So this is sort of the cost that an individual has to pay in order to become resistant. So in this case, while individual B is resistant and A is not, in an environment where there is no antibiotic, they are both going to do just fine. Except for the fact that because B is carrying that mutation, it's going to go a little slower as compared to individual A. And as a result of that, A grows faster. Compared to B. And we have seen that we've been using growth rate as a proxy for fitness anyway. So this is a proxy for fitness.

What that implies is that individual A is fitter than B. So, using this example, we have seen how in one environment individual B is fitter, but if we change the environment, individual A becomes the fitter one. In the context of, if you look at this discussion in the context of this particular mutation, what we are saying is that in this particular environment which contains antibiotics, this mutation is a beneficial mutation. So, the mutation that we are discussing here is beneficial. However, if I change the environmental context, this mutation is deleterious.

So, again, I hope this reemphasizes the fact that the effect of mutations is dependent on environmental and genomic context. All right. Now, we also understand. We need a framework to quantify how much benefit a beneficial mutation confers. So if I quantify, if I ask this question: how much benefit does a beneficial mutation confer?

We have a way to quantify this, and this quantification is an extremely important one in all of evolutionary biology. Let us imagine individual A. Let us go back to individuals A and B. This is individual A. This is individual B. And let us say that their fitnesses This individual has a fitness fA. This individual has a fitness fB. And all of this is because of this one mutation that individual B is carrying in its genome. That's the only difference between these two individuals. Then we want to quantify how much benefit this mutation confers.

And the way to quantify that is via something called a selection coefficient. This is represented by small s, and this is the quantity that is extremely important in all discussions of evolutionary biology. So, the selection coefficient of any mutation is described as fB minus fA divided by fA. So, in the numerator is the change in fitness—the new fitness, the fitness of the mutant minus the fitness of the ancestor—and in the denominator is the fitness of the ancestor. And as we will see, much of the discussion about the fates of mutations depends on what the value of s is for a particular mutation.

$$s = \frac{(f_B - f_A)}{f_A}$$

What should become immediately apparent here is that if this mutation is a beneficial mutation, then if this mutation is a beneficial mutation, then obviously fB is bigger than fA. If fB is bigger than fA, that means the selection coefficient is equal to the numerator being positive. So, the numerator is greater than 0, and the denominator is fitness, which is also greater than 0. So, S in that case is greater than 0.

Again, something to keep A note of this is extremely important: the selection coefficient associated with all beneficial mutations is greater than zero. On the other hand, if this mutation was a deleterious mutation. then fB, by definition, is less than fA because this is a deleterious mutation, in which case s will be a number which is less than 0 in the numerator and greater than 0 in the denominator, and hence s is a number which is less than 0. All deleterious mutations have their selection coefficient less than 0.

All beneficial mutations have their selection coefficient greater than 0. And in the third case, where this is a neutral mutation. Again, by definition, in this case, fB is simply equal to fA, in which case s is just equal to 0 in the numerator and some positive number, which is the ancestor's fitness in the denominator, in which case s is equal to 0. So the thing to note here is that beneficial mutations have a selection coefficient greater than zero. Deleterious mutations have a selection coefficient less than zero.

And for neutral mutations, it's exactly equal to zero. Now, we are going to go a little bit deeper into this discussion. And let's take the case of only beneficial mutations. Imagine that we have two different types of mutants here.

So, this is the ancestor, this is mutant 1, and this is mutant 2. Let us call this ancestor, this is mutant 1, this is mutant 2. Mutant 1 carries this mutation, and mutant 2 carries a distinct mutation. Their fitnesses, when I list them here, are fA, fmutant 1, fmutant 2. Let us imagine that both these mutations are beneficial in nature.

They are beneficial, which implies that fM1 is greater than fA and fM2 is greater than fA. But now what we are saying is that both are beneficial mutations, and hence the fitness of both these genotypes is greater than that of the ancestor. The question arises: does the benefit that is conferred by each of these mutations equal each other? And that we can quantify using the selection coefficient. So now we will write the selection coefficient of each of these two mutations.

The selection coefficient of the first mutation here can be written as s of M1, which is fM1 minus fA divided by fA. The selection coefficient of the second mutation is sM2, which is given by fM2 minus fA divided by fA. Obviously, the value of sM1 and sM2 is going to depend on fM1 and fM2, or the other variable in this expression is fA, and that is the same in both these expressions. What that means is that, depending on the precise value of fM1 and fM2, the selection coefficients are going to be different. Let us just take some numbers here and say fA was 1, fM1 was 1.05, and fM2 was 1.03.

In which case, sM1 comes out to be 0.05 and sM2 comes out to be 0.03. This tells me that the benefit conferred by different beneficial mutations is different. This implies that not all beneficial mutations are the same. Each beneficial mutation will confer a different benefit to the individual, and this will have evolutionary implications.

So when we said that mutations can be classified into three categories—beneficial, deleterious, and neutral— now what we are saying is that within beneficial, not all beneficial mutations are the same because each confers its own value. or increase in fitness. And the same discussion can be had for deleterious mutations, where the only difference will be that now in the case of deleterious mutation, fM1 is less than fA, and fM2 is also less than fA. Each deleterious mutation reduces fitness by a different number.

And this particular quantification is not valid for neutral mutations because all neutral mutations are the same. Because by definition, in neutral mutations, the fitness of a mutant is simply equal to

the fitness of its ancestors. So all mutants have the same fitness as their ancestors. So from the context of change in fitness, all neutral mutations are the same. But the point is that now this

We have branched these mutations into these three categories, but within each category, mutations are not the same. And the study of the differences between these two groups, called beneficial and deleterious mutations, is called the fitness distribution of fitness effects, which we'll continue with in the next lecture. © transcript Emily Beynon