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 17

Features of Mutations: Part 2

Hi. So, in the previous discussion, we looked at four major features of mutation, and in this discussion, we will cover the remaining three. So, the first feature that we are going to talk about, a very well-known feature of mutation, is that They are random with respect to what the organism needs. Now, what do I mean by that? Now, suppose you have a petri dish, and in that petri dish, you have put some culture of, say, some bacteria. Which is, let us assume, sensitive to some antibiotic, a bacteriophage, or whatever. Bacteriophages are viruses that kill bacteria.

So, when you do that, and when you spread it, and when you incubate it for, say, 24 hours or 48 hours, You are going to get something like this, where you will have lots of colonies of bacteria on the plate. Now, suppose you hit this with some antibiotics or some bacteriophage; what will happen? Most of the bacteria will die, as will most of the colonies, but you might end up getting something like this where a few colonies will survive. Now, obviously, these colonies have some kind of mutation that allows them to survive the challenge. Now, here comes the question: what exactly is the mechanism or way in which these mutations have arisen? There are two ways in which this can happen.

One is that the mutations arose after they were challenged with the bacteria in response to them. I am sorry, challenged by the antibiotic or the fudge. So, the mutations are a direct result of the challenge that you have subjected them to. And the other possibility is that

the mutations were already present in the bacterial cells before they were plated. Obviously, they are not in all of them; they were in some of the colonies.

And when you hit them with this challenge, only those that survived had the mutation, and all the others died out. Now, how do you distinguish between these two possibilities? Now, although this looks like a very simple question, this is actually a very big and well-known question in biology. And the primary reason for that is if the mutations arise after the bacteria are challenged. If the organism is being challenged with some environmental factor, then, One can say that adaptation, which is a major feature of most organisms, can be a direct result of mutation. You do not really need natural selection to be the causal reason for adaptation.

So, if the first hypothesis is correct, then mutations will become a major evolutionary force that leads to adaptation. But if that one is not correct, then of course what Darwin said about natural selection will end up getting a lot more credit. So, that is why this was a very big question in biology. And then you know how the distinction between these two hypotheses was made for the first time by these two gentlemen. Max Delbrück and Salvador E. Luria. And what they came up with was this thing called the fluctuation test, a very elegant experiment. and which ultimately led to a Nobel Prize in 1969. So, you can see that you know the importance of figuring out the answer to that particular question. Now, the fluctuation test is a very important technique, and it is still used in certain areas to estimate the mutation rates. However, it is a slightly complicated technique based on the properties of statistical distributions.

Therefore, in the context of this course, we are not going to look at this particular experiment. We will look at another experiment that is equally famous, also leads to a Nobel Prize, but, which is much simpler to explain and is very elegant, logically elegant. And that is the experiment performed by the husband-and-wife team of Joshua and Esther Lederberg. And I am sure all of you have heard about this. This is the famous replica plating experiment.

So, what did they do? So, they started with populations of E. coli that were sensitive to a certain bacteriophage known as T1 phage. So, they took such a clonal population, let it grow, and then at some point from the liquid culture, They plated it on a Petri dish and obtained a bunch of colonies exactly like this. So, this is what they call the original plate on which they plated the E. coli: the master plate.

Of course, on this master plate, the colonies have a certain spatial distribution. Now, they did a very elegant innovation. I am not even sure if the innovation is theirs, but they certainly used it in this particular context. And the innovation was very simple. All they did was take a plastic cylinder and then take a, you know, cloth, a velveteen cloth, basically some kind of rough cloth, and they sterilize the cloth; they sterilize the cylinder. And then they put the cloth on top of the cylinder with some kind of tying, a rubber band, or some kind of thread. So, the main thing is that both the cloth and the cylinder are completely sterile. So, what they did was take this apparatus, and, They took their master plate and simply pressed that velvet cloth on the master plate. What will happen? You have all these colonies; the cloth is rough.

So, from each colony, some number of cells is going to go and stick to the cloth, right? Now, they took this cloth with the bacteria and fresh petri plates containing bacterial medium. And they pressed them carefully on those plates. What will happen? From each point where there were some bacteria, some bacteria are going to be transferred to this cloth, sorry, to this new petri plate. Now, the most important part here is that if you are doing it very carefully, then the pattern of the colonies that you find here, You know that the pattern of the colony on the master plate is going to be replicated in the fresh plate. to which you are pressing the velvetine cylinder.

In other words, you know the positions of these colonies will become the same, and you can do it multiple times. So, these are the so-called replica plates, plates in which you have. Exactly the same colonies replicated onto a different fresh plate. Now, once you have the replica plates, you can ask the question that we posed: how to distinguish between the two hypotheses. What was our first hypothesis? Our first hypothesis was that

the mutations arose after the bacteria were exposed to the phage.

So, what did they do? They had the master plate, created the replica plates, and then added the virus. Now, remember that the mutations arising are a probabilistic event, right? So, not every bacterial colony that is challenged is going to lead to a mutation. Therefore, if it is a probabilistic thing, then you do not really expect. That the same colonies are going to get the mutations again and again. In other words, the mutations are going to be formed in some colonies, but these will be Different colonies for different replica plates, which means that the positions of the colonies are going to be very, very different.

In other words, if you have the mutations that arise after exposure to the virus, The resistant colonies will be at different positions on the plate. Now, what would have happened if the mutations arose prior to the exposure to the pathogen? If they arose prior to the exposure to the phage, then they are already there in the master plate, right? Now, if they are already there on the master plate, then of course they are also already there on our replica plates. And more importantly, because the replica plate is a spatial copy of the master plate, therefore, they are exactly in the same position. And hence, when you add the virus here, what is going to happen? All the resistant colonies are going to be found in the same position on all the replica plates, right? And that was when Lederberg, you know, when the Lederbergs did the experiment. That is exactly what they ended up finding, which told us that the second hypothesis was correct.

All the mutations were already present; the resistant mutations were already there on the master plate. Even before they had been exposed to the fudge. So, the implication of this is that the mutations arise irrespective of what the organism needs. Remember, at the point at which the mutations arose on the master plate, the bacteria had not yet encountered the phages. So, there was no requirement for the phage; they arose independently of that anyway.

And in that context, the mutations are random with respect to fitness, random with

respect to the benefit or harm of the organism. Now, as I said, if this were not true, one could explain adaptation through mutation. In other words, one could say organisms are adapted to their environment. Because each time the environment throws a challenge, the organism comes up with a new mutation. Now, of course, there is a very simple logical problem, you know, with this.

How on earth will the organism know which mutation to come up with? So, there are so many genes that there is no way for the organism to figure out, okay, now I am being challenged with Let us say rifampicin; therefore, I need to mutate the rifampicin gene in my body, which does not really happen. I mean one cannot conceive of a way in which it can happen. But even if it could have happened, the replica plating experiment shows that, actually, it does not. So, adaptation cannot be explained by mutation. And hence, Darwin's theory of adaptation through natural selection gained a lot more credence.

However, the big thing to remember is that although mutations are random with respect to fitness, That does not really mean that they are unbiased in other respects. Actually, one can see that there are many ways in which mutations can be and are actually biased. I will give you just three examples. For example, remember we talked about transition mutations and transversion mutations. So, transition mutations are purine to purine and pyrimidine to pyrimidine whereas transversion mutations are from purine to pyrimidine and from pyrimidine to purine. So, if you remember that square figure we had, you will immediately realize that There are more ways to have transversion mutations than transition mutations. Yet, if you look at the rates, transition mutations are the ones. which are actually much more common than the transversion mutations. Similarly, if you look at our genome, it is not that the genome has an equal probability of mutation at all positions; absolutely not.

There are certain locations on the genome that have much greater mutation rates than other locations. And these locations are known as mutational hotspots. So, the presence of these hotspots straight away tells you that There is some degree of spatial bias in the genome in terms of mutation rate. And finally, in many species, male germline mutation

rates are actually much greater than those of female germline mutation rates. So, as I said, these are just three examples of mutational biases; there are many other kinds of biases that people have uncovered.

The other major feature of mutation is that it is a weak evolutionary force. Now, what do I mean by that? Now, in order to understand the meaning of this, it would be better if we go through an example. So, let us assume that you have a population; we are looking at one particular locus, and, Let us assume that, to begin with, the population only has the allele A2. And at some point, the new allele A1 has arisen in the population. And let us assume that, right at the beginning, the frequency of this A1 allele is 0.001. Now, you can very logically ask if, in a largish population, You have just one mutation freshly arising; how is it that the mutation frequency will be 0.001? Shouldn't it be much smaller than that? And you are absolutely right. In any slightly larger population, if you have a fresh mutation, its frequency will be much, much smaller. However, we are just making a case, okay? It is a theoretical example that we are using. You will realize later that whatever case I am making becomes much, much stronger if we take mutation rates much lower than 0.001. Now, the question that we are asking is, suppose from 0.001, This mutation needs to reach a frequency of 0.1, which is not a very high frequency, but let us say it has to go up to 0.1. How long will this mutation take to reach that frequency if the rates of forward and reverse mutations are 5×10^-5? Now, what is a forward mutation? A forward mutation is A1 converting to A2.

What is reverse mutation? The reverse mutation rate is, sorry, reverse mutation is A1 becoming A2. Oh, sorry, I ended up writing the same thing over here. So, forward mutation is A1 becoming A2, and the reverse mutation is A2 becoming A1. So, this should be 1, and this should be 2. Yeah, A1 is becoming A2, and this is A1. Sorry, this is a typo here. So, anyway, why am I choosing a rate of 5×10^-5? So, it turns out that the rates of mutations roughly vary from 10^-3 in RNA viruses. which you know are the most heavily and highly mutating organisms, to about 10^-10 or 10^-11 or so. That is the whole range, so I have simply taken something in the middle of that, about 10^-5, right? And let us assume that all the other conditions of the Hardy-Weinberg equilibrium are

operative. Now, how will one even solve this kind of issue, this kind of problem? So, in order to do this, what we will do is conduct a very simple simulation. So, this is a link that contains a very nicely done online simulation platform. This is from the same BioInteractive platform that we met earlier. So, what we will do is go there. So, it is free for everybody. So, you can play with it later and have fun with it.

So, let us go there. So, this is the platform. And when you come here, this is where you should go. So, you know, the population genetic explorer start tool. When you do that, it will open a new tab, and on this tab, it has all kinds of information, but you can check those out later. So, it gives you two kinds of simulations: one where you just do individual simulations. one case at a time, and the other where you do replicated simulation, you run multiple replicates at the same time.

So, right now we are going to do individual simulations. Here is the setup for the simulation. The first thing it asks is the population size. So, remember we said that all other conditions of Hardy-Weinberg are operative; therefore, we will set the population size at infinity.

So, we will click it over here. How many generations? We can keep it at 500; it does not really matter. We will change it later in a few minutes, and the starting frequency we mentioned is 0.001. And now I will tell you why we took 0.001, because that is the least count of this particular simulator. If you are doing your own simulation using Excel, Python, or something else, you can obviously keep it at a much lower value. So, here the lowest value is 0.001. Great. Now, we need to make some additional settings. What is it that we are studying? We are studying mutation, so we click on mutation. Here, it gives me the forward mutation rate and the reverse mutation rate. So, I have to do this at 5, and I will put this at 5. This is already at 10^-5, but we can alter it to whatever we want, right? So, in the context of our question, we have put them at, you know, 5 x 10^-5, and then all you need to do is say run simulation.

Remember, our question is how long it takes to go from 0.001 to 0.1? Run the simulation.

500 generations have been completed, and if you just move your cursor over here, it will tell you where the values are. And okay, so you can see that even after 500 generations, it is 0.025. So, it has not really reached 0.1. Okay, so what will we do? We will go up and change the number of generations. So, let us make it the highest that the platform can take, which is 10,000. So, we do that, we come down, and we say, "Run the simulation again." As you can see, it is running pretty fast. And here, now you can see some movement. And let us see where it hits 0.1. So, if you look at it, it hits 0.1 somewhere around 2,329 generations, right? Now, 2,329 generations is actually a very, very big number. So, what this is telling you is that, assuming there is no selective advantage for the mutation, it doesn't provide any benefit or it doesn't cause any harm or anything; even then, with those mutation rates, This is going to take this many generations, which is a very large number. So, this is what I meant when I said that it was a weak evolutionary force. In other words, it will cause a change, but it will cause the change very, very slowly. So, as I said, 10^-5 is still a slightly high number. What happens if I reduce the mutation rate to, let's say, 10^-7 or 10^-9? So, let us do 10^-7 first. We are still keeping the forward and backward rates similar. So, if you have reduced the mutation rates, what do you think? Should it take more time, or should it take less time to reach 0.1? Quickly make a guess and let us run and see it. You can see that even after 10,000 generations, this has only gone up to 0.006. So, obviously, this is taking a lot more time. Yeah, because we are starting from a very low frequency, we are trying to go high. If you reduce the mutation rate, it will take longer. So, by that token, if we increase the mutation rate, it should take a smaller amount of time. So, let us say, let us make it 10^-4 equal for both, and we run the simulation, and there you go. So, you can see that it reaches this, you know, 0.1 much faster; it reaches it about 10 times faster in about 234 generations or so. So, you can see that when the mutation rates are similar, it takes some amount of time. Quite some time, and the smaller the mutation rate is, it takes larger and larger amounts of time. And if we are in biologically relevant zones, which are like, you know, 10^-6, 10^-7 and so on, it is going to take a very, very long time. Now, here, lots of students ask the question of whether I have the same backward and forward mutation rate.

Then why is it that the mutation rate and the allele frequency are changing at all? Because

shouldn't it be the case that the allele frequency, if the forward and backward mutation rates are the same, Then the allele frequency should not change to begin with. That is not really true, because remember when the frequency of the A1 allele is 0.001. Then the frequency of the other one is 0.999. So, at that point, even if the frequency mutation rates are the same, The actual number of alleles that are turning from A1 to A2 is very small because of the forward mutation rate. But the number of alleles that are turning from A2 to A1 is going to be very high. Because even if the mutation rates are the same, they are being multiplied by a larger number, resulting in a larger frequency. And that is why during the initial stage, the rate at which the rare allele is going to pick up is going to be pretty fast and then the closer their frequencies come to one another. So, basically, the closer p and q are to each other, the more the number of alleles that go backward and forward. assuming a same forward-backward mutation rate, that becomes similar. So, I will show that to you in a minute. However, there is something else that is very interesting over here. What is that? You can see that in this particular case, it is stabilizing at some point.

And if you just take it forward, you will realize that it is stabilizing somewhere close to 0.5. In fact, it stabilizes exactly at 0.5, as I will show you by slightly increasing the mutation rate. So, let us say let us make it 10^-3. Both forward and reverse are the same, and when we do that, You can see that it stabilizes, and it stabilizes at 0.5. Now, let us make it something else. Let us go back to 10^-4; let us make it say 9*10^-4. And we run the simulation again; you can again see that it stabilizes at 0.5. So, this property, although I am showing it to you for 10^-4, This property is actually true for any time the forward and backward mutation rates are the same. Now, why is that? In order to understand why that is so, we need to go back to our PowerPoint. And see how exactly the mutation rates behave at equilibrium, or rather how allele frequencies behave at equilibrium. So, let us assume that our two alleles are A1 and A2. The rate at which A1 is becoming A2 is u, and the rate at which A2 is becoming A1 is, let us say, v. so, let us further assume that the initial frequency of this is given by p0 and the initial frequency of that is given by 1-p0. You can also say 1 - p0 = q0. Now, we want to look at Δp. What is the rate at which the value of p, the frequency of the A1 allele, is changing? Remember, that is what we are

interested in. Now, this thing [Δp] = (the rate at which the A1 allele is being created - the rate at which the A1 allele is being destroyed). Now, what is the rate at which the A1 allele is being created? Remember, A2 is becoming A1 at a rate of v, right? So, A1 is being created at a rate v*(1 - p0), correct? And A1 is becoming A2; that is, A1 is becoming not A1 at a rate which is per capita rate u*p0, right? So, Δp = [v*(1 - p0) - u*p0] at this point. So, this is the expression. Now, what happens at equilibrium? So, at equilibrium, Δp = 0; we would not call it p0 anymore because this is at equilibrium, so we will call it just p. So, v - vp - up = 0, right? So, what do I need to do? I will end up taking p as common, rather than (-p) as common. v-p(v + u) = 0 ||| So, this becomes v + u = 0, and I can take p to the other side. So, p*(v+u) = v, which implies that p at equilibrium is equal, so I am just calling it a $\hat{p}$; p at equilibrium = v/(v+u). This is the relationship; this is the general relationship. Now, the special case that we were doing was v = u, right? So, when you have v = u, then $\hat{p}$ = v / (v + v) = v / 2v. So, you just cancel these two out, right? This is why we were obtaining that information: when the forward and the reverse mutation rates were the same, Then the whole population was stabilizing at an equal frequency of the two alleles. Now, what happens if the allele frequencies are not equal? You can figure that out from this particular relationship. So, what are the implications of this? Remember, biologically speaking, the range in which things go is 10^-3 to 10^-10 or (-11) or so. And 10^-3 is only for exceptionally fast RNA viruses and so on. So, mostly things are in the range of 10^-7, minus, you know, 10^-5, (-7), (-9), that zone. In that zone, by itself, mutation will change the allele frequencies very slowly. However, when mutations operate in conjunction with other evolutionary forces, Particularly in selection, mutation becomes a very potent force.

And we are going to look at that as we go along. However, before we get into that, we first have to study selection. But before we study selection, we have to look at one other thing that is very important. So, until now, we have mainly been talking about mutations in the genes, etc. But I also told you that one of the things about EES is that we have now started to realize that There are lots of inheritances happening that are non-genetic, not in our ATGC sequence. So, before we go into other evolutionary forces, we need to take a quick digression. And look at what these other non-genetic inheritances are. And in what

ways they can be effective in shaping our evolutionary trajectories. So, that is something that we are going to do in our next discussion. See you later. Bye.