**Evolutionary Dynamics**
**Supreet Saini**
**Chemical Engineering**
**Indian Institute of Technology Bombay**
**Week 06**
**Lecture 26**

Let us continue our discussion on fitness landscapes of antibiotic resistance. Welcome back, everyone. So, as we saw in the last video, we have two variants of a gene in E. coli. Both these genes are present in nature. The differences between these two variants of the same gene are at five positions.

In the made-up sequences we used, one variant had all As and the other variant had all Gs. And we saw that one variant had an MIC that was really low, and the other variant had an MIC for a particular antibiotic that was a thousand-fold higher. And evolution from the low-MIC to high-MIC variant of the gene can only take place via the acquisition of one mutation at a time. And these one mutations—because there are five options for the acquisition of the first mutation, four for the second one, three for the third one, and so on and so forth— There are 120 different paths available for transitioning from the ancestral variant to the ecological variant, which offers a thousand-fold higher MIC.

And these 120 paths can be classified into one of these two groups. Fitness, where MIC is the proxy for fitness. In green is the path that is permissible in a Darwinian sense. So, the green path is the one that is permissible in what is referred to as a Darwinian sense. And what we mean by that

that natural selection will take this path. Natural selection will take this path. Because along this path, fitness or MIC is monotonically increasing. So we say natural selection will take place because along this path, fitness, which is quantified in terms of MIC, is monotonically increasing. This class of paths is one particular group.

Alternatively, we also have paths which belong to in this blue category. These are paths which are non-permissible in a Darwinian sense. Non-permissible paths. And the reason is because fitness is not increasing. Fitness is not increasing along this path.

As you can see, along this path, the first mutation increases fitness, the second mutation also increases fitness, but the third mutation actually leads to a decrease in fitness, as a result of which we have a local peak here. So fitness is not monotonically increasing, increases. We reach a stage

along this path where fitness increased from behind, but also going forward, it's decreasing. So through this local peak, you move in either direction and fitness is decreasing and hence you are stuck in that local peak because in a Darwinian sense, you cannot decrease fitness. Natural selection will not permit that.

So, and then we asked that in these two groups, how are these 120 paths divided? And what we found out that overwhelming majority of the paths were present in this particular group. There were only 18 paths which were divided. present in this design which was going to be permissible by natural selection. So these 102 paths were the ones which had a local peak in them.

Somewhere along their path there was a local peak and hence these are not permissible. And all of this is a result of epistasis. And this is sort of the first manifestation that we see of epistasis coming into being. Remember that epistasis, if it wasn't there, then we would have irrespective of our starting point, irrespective of the order in which the mutations were acquired, all paths would lead to the highest fitness. The highest fitness in this case, the ecological strain into fitness being quantified as an MIC.

However, because epistasis is playing out, it means that some mutations are beneficial in some contexts and some mutations are not beneficial in some contexts. Let us do this with an example. Let us imagine that we have a DNA sequence and we have two possible mutations on this DNA sequence. Hence, we have four possible genotypes. This is mutation A, also present here, and this is mutation B, also present here.

Let's say that the fitnesses of these genotypes are 1, 1.04, and 0.98. If these are the fitnesses, then looking at these numbers and these mutations should tell us something. What we should do while analyzing this data is compare this ancestral sequence with this mutant. Let's call this ancestor. This is mut one.

This is mut two. And this is mut 1, 2. Comparing the ancestor with mut 2 tells me that the circle mutation, this mutation, when it happens in this background, Because transitioning from here to here is essentially introducing the circle mutation in the black background. This mutation is a deleterious mutation.

Because the delta F associated with it is 0.98 minus 1, which is minus 0.02. Hence, this is a deleterious mutation. But if you look at the effect of the same mutation, but in a different strain now. Now, let us compare mut 1 and mut 2. This is the same mutation occurring, but in a different background.

So, this is the same circle mutation occurring, but occurring in the following background. When in not the plain background this time, but the plain background which is carrying this red mutation. In this case, the delta F associated with the blue background is 1.09 minus 1.04, which is 0.05. And this is the case where a mutation which was originally deleterious has actually become beneficial only because of the presence of this one other mutation.

In this case, this mutation has now become a beneficial mutation. So what we have here is epistasis dictating the nature of the mutation that is taking place in such a context. In one background the mutation was deleterious but presence of just one more mutation makes the same original mutation beneficial transitioning from deleterious. And this is epistasis and this is the reason why we have so many local peaks because some mutations, their genomic context in which they are occurring is extremely important.

And as a result of that, it is often that in a five step path, you will often get stuck in a local maxima and hence not be able to move forward. So in this context, what has epistasis done? To begin the process, we had 120 mutations, not mutations, we had 120 paths. And if we didn't know anything, if there was no epistasis, then perhaps each of these paths was equally likely. This is assuming if there was no epistasis and there was a single global peak and irrespective of starting point, you would converge to that global peak.

However, we know there is epistasis. and figuring out this landscape by Huttle and coworkers told us that of the 120 paths, 102 actually can be ruled out. These could not have been the paths that nature took to go to the fitness peak, and hence the total number of paths that we have to worry about is actually, this is 18, The total number of paths that we have to worry about is actually 18. These are the only relevant paths.

So in some sense, epistasis reduced the number of solutions that was available to me from 120 to only 18. Viewed in such a context, it seems like epistasis has made evolution predictable. more predictable because if it wasn't acting then I had 120 parts to worry about all of them were going to be permissible but because it is acting then I don't have to worry about 120 parts I only have to worry about these 18 parts which lead to increase in fitness however What we also must realize in this case is this was a special case where the end and the starting point were defined to me. This starting point of the evolution experiment, the ancestor lab strain and the ecological strain, they were defined to me.

So the start point and end of this evolutionary trajectory were defined to me. However, when I am thinking of evolution, the start and end are not often defined to me. Very often the question I would ask is that I have this particular variant and if I were to evolve this, what would happen? It is only

because in this study, access was available to this endpoint of evolution that I was able to go back and check, which would not often be available in the context of evolution being predictable or me trying to predict an evolutionary process.

Nevertheless, this was an important study that showed how epistasis constrains adaptive movement, and of all the paths available for a population to take, it only takes a very few. It can only take a very few. Let's now move to the next case of epistasis that we want to discuss—an experimental case of epistasis. This has a little bit of context before we start. So I told you about Stephen Jay Gould asking this rhetorical question: What if we were to evolve the planet again?

What sort of life processes would take place here, and what sort of organisms would inhabit the planet if four and a half billion years were allowed to happen again? And that's, of course, a rhetorical question to which there is really no answer. However, around the same time, a researcher in the United States started the following experiment. The experiment involves evolving E. coli in an environment with low glucose. These E. coli strains growing here in these flasks have low amounts of glucose as the only carbon source available to them.

This experiment was started in the late 1980s. So what you are doing is taking E. coli in a flask, giving it low amounts of glucose, and evolving it in this condition for a long, long, long, long, long time. What the researchers did in this case back in the '80s was not take one flask, but a dozen flasks, all containing the same E. coli at the start of the experiment. So you had an E. coli population in the lab. These were all identical.

And these E. coli were started and were used to seed these populations—these 12 populations in 12 different flasks, under identical conditions. That is another major advantage of doing evolution in the lab: you can maintain the exact same environmental conditions in something as diverse as these 12 different flasks. And in each of these flasks, they were evolved for a long, long time. In fact, we'll spend a lot of time trying to understand major lessons from this experiment because it has become very famous now.

It has a name of its own, and it's called the Long-Term Evolution Experiment. It's also referred to as LTEE. Let me write that again. It's referred to as LTEE. The experiment was started in California.

From there, it moved to Michigan. Now, the person who started it has retired. So, the experiment is now in Texas. Since the 1980s, it's still going on. And recently, it has crossed 80,000 generations of E. coli.

Remember that we discussed that from an evolutionary point of view, it's not the quantum of time. So it's not important that 35 years or so have passed since the start of the experiments, since the

start of this experiment. What is relevant is that, how many generations have been processed. And from the context of this particular experiment, we've processed about 80,000 generations in each of these 12 different independent flasks.

These 12 different flasks are referred to as replicate lines. And In a video coming up soon, we'll discuss how we actually perform these evolution experiments that can go for tens of thousands of generations in a lab setting. So, we have 12 replicate lines. And obviously, this is a study that has been going on for such a long time, giving us remarkable insights into how evolution can take place, what the molecular mechanisms are that drive these evolutionary processes, and so on.

But from the context of today's video, what we want to discuss is just something that was observed in one such line. This is a paper that was published in Science and led by Tim Cooper. And what they did—again, from the context of fitness landscapes—was the following. Imagine that this is the E. coli we started with. After that, we evolved this E. coli in low glucose for several thousand generations.

And at the end of these several thousand generations, I have this evolved E. coli. Both obviously have their DNA. And I sequence the DNA to identify what nucleotide changes have taken place between these two strains at the start of the experiment. And at some time, T. What I note is that in this time that has passed, five mutations have taken place in different parts of the DNA of E. coli. So the number is exactly 5, and this should strike you as important because now we have a mechanism or an experimental setup to test the results that we saw previously from Daniel Hartl's group—what happens to a fitness landscape structure when we have 5 mutations within a gene.

In this context, these were the five mutations that were the first ones to happen in the evolution experiment. They are not confined to one gene. In fact, they are confined to five different genes spread throughout the genome. But we know their identity, we know their location, and we know the genes in which they occur and what sort of changes they drive. So the first thing we want to check is the relative fitness of these.

So we will draw a graph like the one we were drawing for MIC, where this is the start culture, the start bacteria, and this is the five-mutation bacteria. And on the y-axis will again be fitness. And now, in the previous study, we saw that fitness was not directly measured as a growth rate, but it was measured as MIC. That was a proxy for growth. In this case, since they are being evolved in a low-glucose environment for such a long time, their ability to grow

Or their ability to grow fast in low glucose is a measure of fitness. So what is being measured in terms of fitness is the growth rate in low glucose. So obviously, as this population is being

continuously evolved in these low glucose conditions for a really long time, it's going to get better and better adapted to that environment, and beneficial mutations will be acquired. So at the start of the experiment, if this was the fitness, after being propagated for several thousand generations, fitness would increase to a higher level.

And the question that we ask is the same question that Hartle and coworkers asked: now we have the same scenario where we are moving from zero mutations to five mutations. And hence, by the logic that we used in the previous study, there are 120 ways to move from zero to five mutations—how many are adaptive or what we call permissible, and how many are not allowed by natural selection. Remember, in the previous case, we saw that the number of permissible paths was only 18, and non-permissible were 102. If this study also sheds light to the extent that we get similar numbers, then we are beginning to see generality in an evolutionary process.

However, if this study tells us numbers that completely turn them on their head, then we are somewhat lost, as in some cases it gives 102 non-permissible, and in some cases, it gives completely opposite numbers. And we will begin with a discussion of what was seen in the results of these experiments in the next video onwards. Thank you.