Evolutionary Dynamics

Supreet Saini

Chemical Engineering

Indian Institute of Technology Bombay

Week 11

Lecture 51

Hi everyone, welcome back to the course. So, we will begin our discussion of what the results have been from the long-term evolution experiment, and it is impossible to cover everything that has been illuminated in the course of this very long experiment. But let's try to discuss a few preliminaries. So, the evolution, as we discussed, is we have E. coli growing in media that contains low amounts of glucose.

This is the only carbon source present in the environment. The growth period allowed is 24 hours. This is at 37 degrees Celsius, and the flasks are shaken while growth is taking place. So, at the end of this period, the culture density in these flasks is very high, and at that point, we do a 1:100 dilution into a fresh flask containing the same nutrient composition as we started with, and this process continues. So, the first thing is to identify what will happen in the course of this experiment.

When we start our experiment, we begin with an isogenic population. But during the course of growth, many mutations will arise. So, the original genotype will still be present in the flask when growth occurs. But there will also be variants that arise. Different variants will emerge due to the acquisition of mutations during growth.

So these variants will come up. And let us say some of these variants will be growing faster than ancestor variants. growing faster than ancestor, while some of them might be carrying deleterious mutations which are growing slower than ancestor. So, selection that acts on the variation that has been generated in this flask and selection rejects these variants and So these variants are rejected by natural selection.

Whereas these variants will be favored by natural selection. But this is the manifestation of the phenomena that we studied a few videos ago of clonal interference. That in this cartoon here, we have two beneficial genotypes indicated by the green and the blue.

which are competing amongst themselves and eventually given sufficiently long amount of time, one of them will reach fixation. This they are exhibited in a competition between themselves called clonal interference.

And after a sufficiently long time, one of them will win and will reach fixation. Reach fixation. So if we were to quantify fitness in terms of, so through the course of these experiments, this 80,000 odd generations, the proxy of fitness is estimated by the growth rate that these populations exhibit, growth rate in the culture media. That is our estimate of fitness of these populations that are growing in these 12 parallel lines in the LTEE experiment. So obviously as generations move forward, our expectation is, so as generations increase, our expectation is that variants which exhibit higher growth rates are going to be selected.

These variants are eventually going to be able to outcompete the ancestral genotype. So, the ancestral genotype should be eliminated from the population. And the variants which arise due to deleterious mutations are anyway exhibiting growth rates that are lesser than those of the ancestors. So, they will be eliminated by natural selection. So, our null expectation is that as generations increase, the growth rate of the population increases.

Growth rate is my proxy for fitness in this experiment. Growth rate also increases as the number of generations increases. But then, here is a question that we should ask at this point. Suppose I am doing an evolution experiment via serial dilution, and this is going to go for some length of time—let us say a few tens of thousands of generations.

So, let us say this is on the x-axis: time, and this time is measured in the number of generations that have been processed, and on the y-axis is the fitness of the population at any given flask. So, as we are processing these flasks, we start with the ancestor. So, the ancestor, when it reaches its carrying capacity at the end of 24 hours of growth, has a population. So, the mean fitness of that flask is the fitness of the population.

And similarly, as we move forward to the second flask and third flask and so on and so forth, we allow for growth for 24 hours, and at the end of 24 hours, we measure the combined fitness of the population, and that is what is being plotted on the y-axis. So, this is the flask; this is delta T amount of growth. I calculate this is flask number. And this has some mean fitness. And then this undergoes a dilution of one to 200.

And then this is flask number two, and growth takes place for the same delta T amount of time. And at the end of this growth period, this will have a population, and it will also have a mean fitness. And it's not too difficult to see that as this number of flasks increases in this line, fitness should increase because beneficial mutations will keep on accumulating. So as a first exercise, intuitively, let's say at t equal to 0, at the start of the experiment, the mean fitness of the population is indicated by this value called F0. What we do know is that as time moves forward,

fitness should increase. But then what I would like you to do is we will take a 15 to 30-minute pause here. I would like you to sketch what is the way in which fitness increases if I let this experiment run for, let us say, 100,000 generations. In other words, if I let it run for a few decades, just as LTEE has been running for such a long time, and I measure fitness at regular intervals, I will get some data points. If I connect those data points, what sort of trend am I expecting as this experiment is allowed to run for 100,000 generations?

We'll take a 15 second break. And I want you to sketch this on a piece of paper. Whatever you think is consistent with your intuition of this process that is taking place. So let's take a 15, 20 second break now. Okay, so hopefully it does not matter if you are right or wrong, but what I hope to have happened is that all of you have sketched something as to what is your intuitive feel about the process that we are talking about.

And what actually materializes in evolution experiments is that rate of increase of fitness is almost linear to begin with. But then after some time, this rate begins to slow down. And eventually, it keeps getting slower and slower. And this is what the fitness trajectories of most evolution experiment looks like when we perform experiments of the kind that we see below. This is definitely what we see fitness profiles of evolution in the LTEE.

But we know of these profiles from several other examples that evolution adaptive profiles look like this. Fitness increases are much more rapid in the beginning of the experiment and they gradually keep slowing down. So what are the lessons that we know from these experiments? This is not just from LTEE, but countless other experiments that if we were to plot time, which is often measured in number of generations versus fitness. this trajectory has a certain characteristic profile from across different evolution experiments.

And this profile is something like this. In the initial phase, it has a linear increase with time. In the next phase, it begins to slow down. And this slowing-down process continues even when the experiment is propagated for a long, long time. Again, this is something we know from countless lab evolution experiments that people have performed over several decades.

So, this profile here—let us say this is 80,000 generations. This is a representative profile of what has happened in LTEE. But remember that there was not just one line; there were 12 different independent lines. At the start of the experiment, we did not have just one flask being propagated forward in time. That is one line.

We have 12 different exact replicas of the same experiment being propagated. So, what we have plotted here is maybe data for just line number 1, which looks like this. Let us say this is line number 1. So, the first question we ask is: what is the degree of parallelism we observe when these 12 lines, which are subjected to identical selection pressure and whose starting point—in terms of the genotype of the organism they were seeded with—was exactly the same?

So the initial condition is same selection. Pressure that's applied to them is the same. And every experimental procedure that is subjected that these populations are subjected to is identical. Despite that, do we see divergences between these 12 different lines or are the responses of these populations identical to each other? So we will note some characteristic properties.

So this is the first one. The second one that people have noticed and again this is from the context of LTEE and many many other experiments is that when identical experiments these parallel lines are evolved with the same genotype and identical experimental protocol there is also a great degree of parallelism in the response of the populations. So these different colors are responses of different lines from the LTE experiment. So while there are small differences between these different lines but those are quantitative differences of very small magnitude but the overall trajectory and the response of all 12 lines moves as a bundle in a manner which is very very similar to each other. So these are

independent lines from the experiment. While this type of analysis doesn't tell us the nature of mutations that is occurring, so we don't know the changes that are taking place at the genotype level, but when we are testing the phenotype of these populations, that is the growth rate, at least phenotypically, the adaptive response is the adaptive response is are highly parallel so one of the question that comes at this point is that why is this increase in fitness slowing down why is the rate of Increase in fitness. Slowing down with time.

So to understand this, let's think of this process in the following fashion. Let's again go back to the sketch. Fitness initially was this. Fitness after a long time is this. This is time.

This is fitness. This is the ancestor. This is evolved. What we'll do is try and ask: What is the DFE associated with the ancestor? So, DFE, remember, tells us what the available mutations are to this ancestor.

If in the ancestor, this is the fitness effect of a mutation, and this is the frequency. So, what this tells us is that in the ancestor, some beneficial mutations are available, whose probability distribution is given like this. This is the probability distribution of beneficial mutations in the ancestor. But what has happened as this ancestor transitioned through several thousand generations and acquired beneficial mutations is that the availability of beneficial mutations has shrunk for the evolved strain. So if I plot the same data for the evolved strain, the availability will look something like this.

Now, fewer beneficial mutations are available, and the ancestor had this set of beneficial mutations available, which conferred this large benefit. Whereas for the evolved strain, there is no beneficial mutation available that confers a benefit greater than this. So, in the process of acquiring beneficial mutations, what has happened is that the pool of beneficial mutations available to the evolved strain has now shrunk. As a result of this, it is more difficult—the probability of acquiring a beneficial mutation in the evolved strain is smaller compared to the ancestor because the pool of beneficial mutations has now shrunk. And as a result of this, this

It gives this slope of the curve, which was very rapid to begin with, becomes increasingly flat, and so on as we move forward in time. So, I hope this makes sense. Another way to think about this could be to draw a pie for the ancestor and ask: what is the fraction of all mutations in the ancestor that were neutral? What is the fraction that was deleterious, and what is the fraction that was beneficial? So maybe the data for the ancestor was 45%, 40%, and 15%.

But now, in transitioning from ancestor to evolved, several beneficial mutations have already been acquired. So when it comes to the evolved strain, this happens. The pie has changed its shape to maybe this: neutral mutations are still 45%, deleterious mutations are, let’s say, 52%, and beneficial mutations are only 3% of all that could possibly be introduced into this genome. As a result, while mutations are happening randomly in the genome of this evolving population, it is now increasingly likely that they will be deleterious, and sampling these beneficial mutations becomes progressively harder. This led to an interesting development in the context of the LTEE.

What was seen at the start of the experiment was that the ancestral individuals were growing and acquiring mutations at a rate mu. This mu for E. coli, which we have been talking about, is 10 to the power of minus 3 per cell. So, these errors happen as a result of this underlying mutation rate. Now, what happens is that as populations adapt to this environment, we will represent this process in the context of a fitness landscape. I hope all of you remember the concept of a landscape.

Let's say this is the adaptive peak. Let me get rid of this. So, let's say that the population is starting from some point on this landscape, and it's moving up for some time. However, soon, effects like this start to happen. Now, what is happening is that because some beneficial mutations have already occurred and become fixed in the population, access to newer beneficial mutations is very hard.

So, what strategy can cells take to increase access to beneficial mutations? So, let me just draw this: the pie changed from 15% beneficial mutations to 3% beneficial. So, as a result of this scarcity of available beneficial mutations, the adaptive process is going to slow down. And this error rate, if you think about it, is quite small because it only means one error in a thousand divisions.

So what has been seen in roughly 6 lines out of 12 in the LTE experiments is that the mutation rate in these 6 populations has increased considerably. And how will this help? So, imagine in a flask, I have two variants. One which has a high mutation rate and the other one which has a low mutation rate. So, let us say this individual is a low mutation rate individual.

Now, what has to happen and let us say that the fitness of both of these is F naught. So, as far as their fitnesses are concerned they are identical at this moment the only difference between the two is that the green individuals do not pick up mutations very fast whereas the red individuals make errors much faster as compared to the green ones so in the low mutation rate divisions are going to take place errors are going to be very rare so when an error takes place But more often than not, error will not take place. Most replications are just error free.

So only once in a while errors will take place. And in case this error does not increase fitness, let's say this was a deleterious mutation. In case this error does not increase fitness, then this is going to be rejected by natural selection. And as a result, the search for a beneficial mutation continues where errors are happening at a very, very small rate. On the other hand, if this population in the red population errors are happening at a much higher rate.

and so on and so forth. Let us say this error was deleterious. So, this was rejected by natural selection. So, this does not matter, but because and hence the search for a beneficial mutant continues, but because errors are happening much more frequently because this population has a much higher mutation rate as compared to this population. Because of that,

The sampling of different mutations that is being done by the red population is much higher as compared to how the green population is able to sample mutations. And as a result, what's likely is that the red population will soon find a mutation, finds a mutation that increases growth rate. And as a result, now that this individual has come into the population, it will outcompete both its ancestor, but also the green competitor in the flask. And very soon, this will move to a situation where only progeny of the high mutating, high fitness individuals are there in the population.

And as a result, population's mutation rate will increase because all individuals with lower mutation rates have been eliminated from the population. Hence, change in mutation rate sometimes gets selected for not because it increases the fitness. Change in mutation rate doesn't increase fitness, but it makes the access of mutations which will increase fitness easier. And hence, increased mutation rate is sometimes selected for. And this has happened several times in almost half of the 12 lines evolving in LTEE.

We will continue this discussion in the next video. Thank you.