Evolutionary Dynamics

Supreet Saini

Chemical Engineering

Indian Institute of Technology Bombay

Week 9

Lecture 42

Hi everyone, welcome to the next video of the course. So, we will continue our discussion of the two regimes in which evolution can take place in a microbial population, be it in a batch culture of a flask or in a continuous culture of a chemostat. So, what we have to figure out is which of these two regimes is a more representative regime of how evolution takes place in a microbial population. In the first one, we had this scenario where if we plotted fitness versus the number of individuals, evolution took place as follows: we had all the population at one particular fitness, which we called F naught, and after some time, one individual arose which was of higher fitness, so I had this one individual.

And the waiting time associated with these new beneficial mutations to arrive and survive drift is sufficiently large that I have to wait for a mutation to come. And this happens over a relatively longer frame of time. Once that happens, this genotype increases in frequency, and this genotype starts to decrease in frequency. As a result of that, when I wait for a certain amount of time, the resulting picture I get is that the green genotype is completely eliminated from the population, and what I end up getting We'll just assume F naught equals 1.

So, the new fitness of everyone in the population is now 1 plus S naught, where S naught is the selection coefficient associated with this beneficial mutation. This here happens relatively fast as compared to the waiting time associated with the arrival of a new mutation. And we said that, of course, this is going to repeat itself. Eventually, there will be another mutation that comes up, and the same process will keep on repeating itself. This was regime 1.

And this happened in the case where tau wait was much bigger than tau fix. And in an alternative representation of what is happening in this case, what we saw was that if we

represent this process in this way, where x-axis now is time and y-axis is number of individuals, then at t equal to 0, so this varies from 0 to n. At t equal to 0, I'm starting with an isogenic population, which is represented by this green population at t equal to 0. There was no red individual at t equal to 0. So, I start with an isogenic population and beneficial mutations keep coming up in this population, but they don't survive drift.

They keep getting lost. Eventually, there will be a beneficial mutation that survives drift and then it goes to fixation relatively quickly. And we saw that this waiting time is tau wait. And this time, once it escapes drift and reaches fixation, this time is tau fix. And as we can see, tau wave, it is significantly larger as compared to tau fix.

And the same process keeps on repeating itself that now we wait for the next beneficial mutation. But the wait time associated with this large, eventually it will come and reach fixation. And the same process keeps on repeating itself. And the big point to note here was that if I were to take a snapshot of this population at any given time, suppose I chose a random time t dash, if I looked at this population at time t dash, it will appear to me that the entire population is made up of only one genotype. Suppose I sample this population at another time, again it is just one genotype and again and again and so on and so forth.

It is only going to be a very small fraction of time for which there is more than one genotype available in the population. Suppose I happen to sample at a time somewhere in this window, and at this point, I can see that there is one genotype and another genotype. But most of the time, there is only one genotype present in the population. So, this was regime 2. As against this, we have—I am sorry—this was regime 1.

As against this, we have regime 2, in which these time constants—their relative strengths—are flipped. So, now we have a scenario where tau weight is much less than tau fix. So, the waiting time is very short, which means the supply of new beneficial mutations is actually rapid, and the time it takes for each beneficial mutation to go to fixation is longer. So, while it is going to fixation, newer mutations will happen because tau weight is so much smaller than tau fix. In that scenario, the first representation changes as follows: we have this, we have—

all the individuals at a certain fitness F naught, which is one at the start of the experiment. And then what will happen is that another beneficial mutation arises, gets established, and this begins to rise, while this begins to decrease. But unlike in regime 1, well before this reaches fixation—so suppose this is at this frequency, this has come down to this

frequency—another mutation happens here, which gives rise to an individual that escapes drift and is established at 1 plus 2s0. The fitness of red individuals is obviously 1 plus s0.

And all of this is under the assumption that all beneficial mutations confer the same benefit of S naught. But we can relax that assumption, and the picture will still look something like this—a bit more messy, but the essential principles remain the same. As a result of this, because this keeps on happening, as this increases in frequency, another beneficial mutation will occur. What we will end up getting at a given point in time is the following picture. We will have these genotypes, something like this.

These genotypes are older ones in the population, which are going extinct. And these genotypes—highly fit genotypes—they have just arisen in the population and are hence currently small in their numbers, but their frequencies are increasing with time. So that's the kind of picture that you get. And the point here is that this is a snapshot of a population at a given moment in time. And as we can see, at this instant of time, multiple genotypes coexist.

So, it is not just one or two genotypes, as we have shown in this picture—there are seven or eight. And this scenario, in this other representation, changes to the following. Again, at t equal to zero, we start with an isogenic population, but eventually, now I will not show these mutations which are going—which are not able to escape drift. So, a beneficial mutation arises and increases in frequency. But as it is increasing in frequency,

Another beneficial mutation arises. So, we have three genotypes now. The ancestral one here, this blue genotype here, but also now this new genotype. So, this is going on, but as this is going on, another beneficial mutation arises and that increases in frequency. So, for instance, now let us just do a little bit more.

Let us say we are talking about this instant in time. At this instant, what are the genotypes that are present? At this instant, this much fraction of individuals is the green individuals. This much fraction is the red individuals. Above that, this little fraction is these individuals; above that, this fraction is the blue individuals. This is just the blue line continuing, and the last little bit

This little bit is the ancestral genotype that we started with, which still hasn't gone extinct. But as you can see, these are all increasing. So eventually, we will reach a stage where, at this point, the ancestral genotype is completely gone from the culture. So, in

this scenario here, these five genotypes are coexisting with each other. But that will change as time moves forward in this setup.

So the question that we have is which of these two regimes is a more relevant representation of what is happening in a flask. So we will do some calculations to give us some sense of what is happening. So let us imagine that we have a population size which is n. And what that means is that there are n individuals in this population. And for the sake of simplicity, let us imagine that each of these divides in one generation.

And of the two progeny, one is chosen and the other one goes out of the chemostat because the population size remains constant with time in a chemostat. So, if we retain everyone, then the population size becomes 2N. That is beyond the carrying capacity of the chemostat. So, this is generation IA. And as we grow from generation I to generation I plus 1, every individual divides once, and of the two, one is retained and one goes out with the exit stream.

It does not have to be that simple, but this just helps us think about what is going on. So, what that means is that—let us forget about the one that is going out from the chemostat. So, this one goes out; we do not worry about it. What that means is that every parent gives rise to one progeny via division, and the other progeny is simply washed away by the exit stream. So, how many cell divisions took place to give rise to the I plus 1th generation starting from the Ith generation?

Every individual came as a result of one cell division process. So, that means number of divisions that took place is equal to n, right? Because every individual in the I plus 1th generation is coming as a result of a cell division process. So, the number of divisions that took place to make up the composition of the I plus 1th generation is simply n. Now, we also know that the number of mutations is dependent on the mutation rate of the species. So, we need to know about mutation rate.

And as we discussed earlier in the course that mutation rate is simply described as mu and one measure of this is number of errors per cell division. And for E. coli this is around 10 power minus 3. That is why we say that of 1000 cell divisions one cell division will have an error. However, what we have to note is that this mutation rate comprises of all sorts of mutations whose effects on fitness can be very different. So, this mutation rate is the error associated with is the rate associated with all types of errors that are happening in the culture.

So, this includes mu-D which is the mutation rate of deleterious mutations. This is also inclusive of mu-N, which is the mutation rate of neutral mutations. And this also includes mu-B, which is mutation rate of beneficial mutations. In fact, as we can see, mu is simply the sum of all three, mu deleterious, mu neutral and mu beneficial. Again, remember that these are highly environmental context dependent.

So, there might be a particular mutation in E. coli, let us say this circle mutation. This circle mutation in environment A is a beneficial mutation, but the same mutation in environment B can be a deleterious mutation. So, its contribution will count towards mu B if I am if I am doing this analysis in environment A, but its contribution will count towards mu D if I am doing the same analysis in environment B. So, these these things are context dependent and the environment has to be taken into account when we are doing this sort of analysis. But what you should realize that the mutation rate that we've been talking about so far is an all-encompassing mutation rate which takes into account all types of mutations depending on their effect on fitness.

For the context of what we discussed just now, for the context of discussion of what we have here, we are only interested in mutations which increase fitness. We are not talking of mutations. Of course, in this culture, deleterious mutations are also happening. Those individuals will go here. But we know that selection gets rid of them.

Hence, we are not worrying about these neutral mutants which are constantly, I am sorry, we are not taking into account these deleterious mutations which are constantly occurring in the population. We are also not worrying about neutral mutations, which will give you an individual of the same fitness. So it just adds to the current fitness, but is a different genotype. That's a neutral mutation. And we are ignoring that also simply because

The idea is that the number of neutral mutants created via this mutational process is going to be so small as compared to the entire population that I have at that fitness level that their contribution can be neglected. Hence, the only type of mutational events that I am worried about in this entire analysis is associated with beneficial mutations. As a result of that, the mutation rate that I am interested in is only mu beneficial. So when I'm doing this analysis, I am actually only interested in mu benefit. Now clearly, since mu is a sum of all three, we know that mu beneficial is less than total mu.

Because it's only a fraction of all mutations that are beneficial. So if mu is 10 power minus three, We know that mu B is less than 10 power minus 3, but we do not know by how much is it less than. If all beneficial mutations are 10 percent of all mutations, then

mu B is simply 10 power minus 4. But if beneficial mutations are actually only 1 percent of all mutations that are occurring, then mu B in that case is 10 power minus 5.

And this might also be dependent on environment. In some environments, there will be lots of beneficial mutations that are available that an organism can acquire and increase fitness. In those environments, mu B will be higher. But there might be environments where the organism is extremely well adapted. And there is not a lot of beneficial mutations remaining for the organism to acquire to increase fitness.

To give you a little example of this, Saccharomyces cerevisiae yeast. So this is the standard cartoon that we use in literature to draw yeast. This is Saccharomyces cerevisiae. This is the yeast that's used for baking. When in its evolutionary history, it's extremely well adapted to glucose.

adapted to glucose. What that means is that if I evolve this further in an environment which contains glucose, then the increase in fitness will be small. is small and one of the reason is that mu B is actually small in that environment. On the other hand, if I evolve this organism in another sugar called galactose, the yeast is not that well adapted to grow in galactose. And as a result, mu B is higher.

Fitness increases are much more in this case as compared to this case. So what we are saying here is that this all we know is that mu B is less than mu. But how much is it less by is dependent on the organism, the environment, the genotype that exists and so on and so forth. So at this point, we only take into account that mu B. At this point, we want to make note of two things. One, that in our analysis here, we are interested in mu B and not mu.

And second, all we know of mu B is that it's less than mu. Maybe it's 10 times less than mu. Maybe it's 100 times less than mu. But those precise numbers are context-dependent. All right.

So if I have In one generation, I had N divisions take place. The mutation rate of beneficial mutations was mu B. This is the chance that one cell division leads to progeny which is beneficial, carrying a beneficial mutation. Then the total number of beneficial mutations in this generation is simply equal to N mu B. This is the supply of beneficial mutations.

But we know that of all the beneficial mutations arising, most will be lost because of drift. So, the probability that they survive drift, as we know, is simply equal to S naught.

Where S naught is the selective advantage that each of these mutations confers to the population. Obviously, mutations that confer a higher S advantage have a greater chance of escaping drift. But as we saw in this particular analysis that we are doing, we are only concerned with beneficial mutations which confer the same selective advantage of S naught.

So, So, what that means is that of all these occurring, this is the probability of survival. Hence, this is total number of beneficial mutations in one generation. Because n is the number of divisions that took place in one generation, mu is the beneficial mutation rate. And hence, their product is the total number of beneficial mutations that occur in one generation.

So, that means that number of beneficial mutations that survive drift in one generation is simply equal to N mu B times S naught. this is the number of beneficial mutations that occur in one generation, and this is the probability that each one of them survives drift. So, if in one generation, number of beneficial mutations that are occurring is N mu S naught, then the question is that how long does it take, by how long we mean number of generations, for one beneficial mutation to escape drift. To escape drift.

And if you read this carefully, what we are asking is simply tau wait. And this is just unitary method where we say that N mu S naught mutations escape drift in one generation. So, in how many generations to one beneficial mutation, does one beneficial mutation escape drift? That is simply equal to 1 upon N mu S naught generations. So, this is simply equal to tau wait.

And you can get some sense of, some relative sense of the magnitudes involved here. How long is tau weight in generations? And we can see a typical microbial population will be of the order of, let's say, 10 to the power 8, maybe 10 to the power 9, maybe slightly larger than that. Mutation rate, this is not the overall mutation rate, but this is mu B. So let's say. The supply of beneficial mutations is only 1% of all mutations that are occurring, in which case this will be 0.01 times the overall mu, but we know this mu is 10 to the power minus 3.

So we are looking at 10 to the power minus 5. And S naught is the benefit that is conferred by a beneficial mutation. We know this is of the order of 3%, 4%. That is a good beneficial mutation. But let us say this is 0.02.

So this is 2 into 10 to the power minus 2. So, we have some sense of how long tau weight will be, at least an order of magnitude estimate for a population of this size. This is simply equal to 10 to the power 8. Let us do it for 10 to the power 8 times 10 to the power minus 5 and then 2 into 10 to the power minus 2. These two are 10 to the power minus 7.

This is 10 to the power of 8. So, this is 1 over 20. which is 0.05. So what that means—what this number is telling us—is that the waiting time for a beneficial mutation to arise and escape drift is actually much less than one generation. So in one generation, multiple beneficial mutations arise and are able to escape drift. What does increasing the population size do? Let us say, let us calculate tau weight, but now increase the population size from 10 to the power of 8 to 10 to the power of 9, which is still not very large.

Mu B was 10 to the power of minus 5, and the mutation rate was 2 times 10 to the power of minus 2. This just becomes—this is 10 to the power of minus 7—this is 9 becomes 2. So, I have 1 over 200. which is 0.005 generations. So, as we see, when we were talking about these two regimes, we were talking about tau weight and tau fix.

So, the waiting time associated with a beneficial mutation arising and escaping drift turns out to be much less than one generation. That's what these numbers are telling us. In the next video, we'll start our discussion with how to estimate tau fix and what that eventually tells us about the likelihood of operating a microbial population in regime 1 or regime 2. We'll continue with that in the next lecture. Thank you.