Evolutionary Dynamics
Supreet Saini
Chemical Engineering
Indian Institute of Technology Bombay
Week 02
Lecture 11

Thank you. Hi everyone, welcome back to the next video of the course. So we'll keep moving ahead and understand in this video how diversity is generated in a prokaryotic population where reproduction is asexual. So the first thing we should realize is that we have an individual which grows under appropriate conditions, and this individual then divides to give two individuals. And this type of inheritance, where one individual gives rise to two, is called uniparental inheritance.

Uniparental inheritance because what an individual has acquired has only come from one parent and not two parents, such as the case with complex organisms like human beings. So, in such a context, what we were discussing in the last video was that if we have a microbial population where at t equal to 0 the number of individuals is let us say equal to N naught, and we have this population in a growth flask which contains appropriate nutrients to support cell growth and cell division. If we keep this population under these conditions and we let some time pass, then at a future time T equal to T naught, the number of individuals would have increased to Let us say NT.

And we need to understand a few properties associated with this process. So the first one we looked at last time is how many divisions took place in this time. And that number is simply equal to the number of divisions in t0 time, which is equal to nt minus n0. This number of divisions is equal to this expression because of the observation that every cell division increases the count of the population by exactly one. So one cell division implies an increase in population by exactly one. Hence, the increase in population that we see here in the case that we are looking at is simply nt minus n naught, and hence the number of divisions that need to take place to facilitate an increase from nt to reach n naught is simply equal to nt minus n naught.

That's the first thing that we note about growth taking place in this flask. The second thing that we want to note is how many generations passed in this time. How many generations passed in T0 time? And to calculate that, what we should realize is that, let us do this on the next slide, that we started at t equal to 0, the number was n naught, and we let growth take place for a while, t is equal to t naught, and the number of individuals in the flask is n t. So, this transition has happened in this time, and we want to understand that number. We now know that the number of cell divisions in this process that took place over t naught time is simply equal to the number of cells at the end minus the number of cells at the beginning of the process.

And what we are interested in is how many generations happened in this time. So, one way to think about this is that at the start, the number of individuals was simply equal to N naught. After one generation, the number of individuals. So, after one generation, what has happened is that if there were N naught cells, after one generation, every one of these cells has divided exactly once.

So, instead of N naught, we have two times N naught cells after exactly one generation. So, after one generation, the number of individuals in the flask is going to be 2 times N naught. Similarly, let us do one more: after two generations, the number of individuals will be 2 times 2 times N naught, which is 4 N naught or simply 2 to the power of 2 N naught because now each one of these, so we started with N naught, after one generation we had 2 N naught, and now after another generation, every one of these 2 N naught individuals is going to divide, and we are going to end up with 4 N naught individuals. What you should realize at this point is that if we keep doing this process over and over again, we will get this relationship: the number of cells after K generations is simply equal to the number of individuals after K generations, which is simply equal to 2 to the power of K N naught.

Because for every generation, we simply multiply the number of existing individuals by 2. Because in one generation, every individual divides exactly once. Hence, at the end of K generations, we have this many individuals. And obviously, in the scenario that we are discussing here, it's the value of K that we are interested in, that is, what would be the number of generations that took place which increased the population size from N0 to Nt. And to find that value of K, we simply have to equate these two numbers, and we get 2 to the power of K times N naught is equal to Nt.

So we solve this since Nt is given to us and N naught is known to us. What we have to do is find for what value of K this equation is true, and we do this simply by taking the log. We can take N naught to the other side, and we get 2 to the power of K is equal to Nt divided by N naught. Or if we take log to the base two, K is simply equal to log base two of Nt divided by N naught. That's the number of generations that happened while this expansion of population took place as the population changed from size N naught to size Nt. What that means is that if a population increased from, let's say, size 100 to size 400, Then the number of generations K is equal to log base 2 of 400 divided by 100, which is simply log base 2 of 4, which is just log base 2 of 2 to the power of 2.

The 2 comes outside, and then log base 2 of 2 is simply 1. So, this is just two generations. So this gives us mathematically, but we can simply see here that if you were to think of 100 and 400, after one generation, everybody divides, and this becomes 200. So we have 200 individuals. And in the very next generation, each one of these 200 individuals divides, and we get to a population size of 400.

So this is generation number one, generation number two. So exactly two generations is what it takes for this population to become four times larger. And we can keep on doing these experiments, keep on studying these numbers of the kind to understand how many generations it took. So that's about how many generations it was, how many generations it took for the expansion of the population to take place, or how many cell divisions happened, and so on and so forth. One key thing that needs to happen for cell division to take place is that if you think of an E. coli, while we have been looking at the external size and number of the bacteria, something needs to happen internally inside the bacteria before it's ready for division.

And one of the key requirements of cell division is that the DNA must be copied. So let's imagine that this is an E. coli cell. This is its DNA. And after division, this cell is going to give me two individuals. But each one of these individuals is also going to carry the DNA.

So what that means is that as part of the growth process, DNA must be copied. And that makes sense. And that way, there are two copies of DNA, and each of the two progenies inherits one such copy. E. coli's genome that needs to be copied is different; strains of E. coli have different lengths, but it's 4 to 5 into 10 power 6 nucleotides. Division is not division; replication of DNA is done by a machine called DNA polymerase.

It should strike you immediately that the process of transcription was done by a machine called RNA polymerase. It was a machine that synthesized the polymer called RNA. DNA polymerase is a machine that is responsible for the synthesis of a polymer called DNA. So, division is done by DNA polymerase. Since this is essentially a cellular machine, and no machine in any context works with 100% efficiency.

In the context of DNA replication, what that means is that as the DNA of an individual is copied into two, as the cell prepares for cell division such that each of the two progenies inherits one copy of the DNA, this replication process will not be 100% accurate. And some errors will be made in the process of DNA replication. What that means is that we are talking of literally millions of nucleotides like ATGGs and so on and so forth. This is circular DNA. So it circles back onto itself.

And for cell division to take place, all of this has to be copied. So let's say the copy will also be ATGG and so on and so forth. And each of the two progenies gets a copy. However, with some rate, This machine that is responsible for copying the DNA and making the additional copy makes errors.

And those errors can be of different types. But to take an example, what might happen is that while DNA is being copied, instead of a second G here, it inserts a C. What that means, and let's say this

is the only error that takes place in the entire 4 to 5 million bases that are being copied in the E. coli genome. What that means is that between the two progeny, this one has this mutation where it's a C here, whereas in the original genotype, this was a G. So this is a mutation that has happened, and this happens because DNA polymerase has a finite efficiency. And by efficiency, we want to quantify what the rates are at which DNA polymerase is making errors while it's copying DNA.

So this is an example of a mutation. And this type of mutation, where a nucleotide is simply written as NTD, changes. In this case, the change is negative. From G to C is called a single nucleotide polymorphism. And we'll study loads of examples in experimental evolution where such kinds of mutations facilitate adaptation and evolution.

And this, in short, is simply written as SMPs, and they're pronounced as SNPs. So whenever we use the term SNP, what we mean is that a single nucleotide change has happened, which will lead to a change in some manifestation of the cell. So the question then is, this finite efficiency, what is the rate at which these mistakes are happening? What is the rate at which these mistakes What you should realize is that this mistake that the DNA polymerase is making is responsible for generating diversity.

The more mistakes that the polymerase makes, the more different types of individuals you are going to get in the population. So this DNA polymerase error rate is responsible for generating diversity. For generating diversity. So what we're going to try and do now is try and understand that in the flask example where the population went from N0 to NT, what is the sort of diversity that is generated purely because of these errors taking place because of DNA polymerase not working at 100% efficiency. So in E. coli, the polymerase makes an error.

The error rate in the polymerase is of the order of 10 to the power minus 3 per cell per generation. What that means is that if I have an individual cell, let us say I have an individual cell. This divides to give two progeny. And each one has the DNA, one copy of DNA. Then this number, of course, is just 0.001.

This is roughly equal to this, 0.001. What this means is that there is a 0.999 chance that the progeny's DNA is exactly the same as the parent DNA. And remember, each of these divisions involves copying 4 to 5 into 10 power 6 nucleotides. Let us just call it 5 into 10 power 6 nucleotides that have to be copied for one cell division to take place.

And even with copying that number of nucleotides, the chance that the progeny DNA is going to be exactly the same as the parent DNA is almost 0.999. So this is obviously equal to a 99.9% chance that there is no error. So DNA polymerase is a machine that operates with a very, very high

efficiency. However, there is this 0.001 or 0.1% chance that an error is made. And if that happens, what that means is that the progeny DNA is not the same as the parent DNA.

One of the individuals is carrying a mutation. And this error rate has a symbol in all of evolutionary biology, and that is mu. So mu is the mutation rate or the error rate associated with DNA polymerase. Hope that makes sense. It seems like DNA replication is an extremely faithful process.

There are hardly any errors that take place. Even in these 0.1% cases where errors are made, the error is simply one error. So this might be a G here, which might change to a T here, and so on and so forth. So errors are extremely rare when it comes to DNA. DNA polymerase working in an organism.

The way we are going to look at this, the way we are going to look at these numbers is the following. Instead of looking at this in probability terms, we'll see, we'll look at it in a slightly different context. That of the 1000 divisions Taking place in any flask, test tube, pond, or any environmental context, we can divide these 1000 into 2. In 999 such cell divisions, there is no error.

In 1, however, there is an error. And this has interesting implications, as we will see in the next slide. So, let us see how big or small this number is by going back to our flask example of population going from N0 to Nt. So, to make our analysis a bit easier, let us just imagine that when we started the experiment, there was only one individual. So, N naught was equal to 1.

We let it go for a long time, and N t is very large. So, this is a long time. 24 hours is long enough time that all the media is just completely used up of the carbon and other resources that are present. And it's just cells with no additional resources to further increase the number of viable cells present in this flask. So, one such popular media that is used for E. coli is called LB growth media.

And this is when E. coli is grown in LB at 37 degrees Celsius with shaking. These are conditions that are very favorable for the growth of bacteria. And typical densities that the bacterial population can reach under these conditions are 10 to the power of 9 cells per milliliter of this flask. So let us imagine that the liquid volume that we started this culture with was 100 mL. And at the beginning of the experiment, there was only one individual in this flask.

So at the end of this long time that we are letting growth take place, the number of individuals in the flask is simply equal to the volume of the flask times the number of individuals. So we know that the volume of the liquid culture that we have provided for the bacterial population to grow into is 100 mL. And we also know that each one of these 100 mL supports 10 to the power of 9 bacteria. So this is 10 to the power of 9 individuals per mL. And when we multiply this, this cancels out, and we get that the number of individuals here is 10 to the power of 11.

So that is the size of the population that we have at the end of this period of growth. Now, we should look back and see, when we look at these numbers, how many cell divisions took place? How many cell divisions? And we know this from before that if n_t is the final number of individuals, n_0 is the starting number of individuals, and the number of divisions is just n_t minus n_0, which is just equal to 10 to the power of 11 minus 1. If you do this subtraction, you should convince yourself that this is basically 1 followed by 11 zeros: 7, 8, 9, 10, 11 minus 1. So, what you get is 9, 9, 9 all the way to 9.

This is simply equal to 0.999 going all the way into 10 to the power of 11. Now, for all practical purposes, this is essentially equal to 10 to the power of 11. This number is just 10 to the power of 11 minus one, which is this to be exact, but we can easily approximate it as simply 10 to the power of 11. So that is the number of cell divisions that took place in this flask. But we know that an error is made.

And one error is made every 1,000 divisions. Every 1,000 divisions, 999 are accurate, and one error is made. So an error is made, one error is made every 10 to the power of minus 3 divisions. The number of divisions that have taken place is this many. So what you should realize is that if we were to quantify the number of errors that are made in this flask, it is simply equal to the error rate times the number of divisions.

Which is equal to 10 to the power of minus 3 into 10 to the power of 11, which is just equal to 10 to the power of 8. So in this one flask that you grew overnight, there are 10 to the power of 3 mutations that took place. This is the number of mutations that happened in the flask. What you should realize at this point is that this is an extremely large number. Remember that the genome of the organism, this individual organism and its DNA, the DNA itself is only of the size 5 into 10 to the power of 6.

At every one of these 5 into 10 power 6 positions, there are only 3 mutations that are possible. The mutations that are possible are that if there is an A, it can mutate to T, it can mutate to G, or it can mutate to C. 999 times it will stay as A. So, these are the only 3 mutations that are possible at every one of these 5 into 10 power 6 positions. That means the total number of mutations that are possible is simply equal to 3 into 5 into 10 power 6.

This 5 into 10 power 6 is the size of the genome, and this 3 is the number of mutations possible at any one point. So, you could be any one of the four nucleotides, and you could acquire a mutation and become any one of the three remaining nucleotides. So, this is equal to 1.5 into 10 power 7. This is the total number of SNPs which are possible. The number of mutations that happened in

that one flask that we grew in the lab is much greater, approximately seven times greater than the total number of mutations that are possible in an E. coli genome.

And we'll continue our discussion about the implications of this in the next video. © transcript Emily Beynon