Evolutionary Dynamics
Supreet Saini
Chemical Engineering
Indian Institute of Technology Bombay
Week 02
Lecture 10

Hi, welcome back, everyone. So, with everything that we have done, we are sort of primed to start our discussion of how evolutionary change can take place in populations. And what we'll do in this video is sort of look at prokaryotic populations and understand how diversity is actually brought about in these systems. So, let's imagine that we have an isogenic population. This is a word that will keep occurring throughout the course.

This is an isogenic population. Isogenic means that every individual in the population has the exact same DNA sequence. So, let's consider two, let's consider a population of bacteria. These can be very, very large, as we'll see in this video. We'll sort of try and get a sense of what the typical numbers of a bacterial population are.

But let's say we have bacteria 1, 2, 3, I, J, N. So, this is a population of size N. The size of a population, as we will also see, is an extremely important evolutionary variable. So, in an isogenic population, let's imagine that we take individual number I, individual number J, and we have there, these are bacteria. So, let's say this is I, this is J. Let's also represent their DNA sequences. And

So these are the two individuals. And these are populations of size N of, let's say, E. coli bacteria. Now, one thing is that because this is the same species, the same strain of E. coli bacteria, the genome length is the same in both. The genome length is L. And if we try to match, if we begin to match the nucleotides of these two sequences, let us say we match that there is an A here. We find that there is an A here.

Next to it is T. Next to it here also is T, and so on and so forth. If we keep going and these two sequences match exactly with each other as we go around the sequence, then these two individuals are said to be isogenic. On the other hand, let us imagine that in this genome length L, these two sequences match at L-1 positions. But there was this one position where there was a G here, but in this particular organism, this particular individual, there was a T here. Hence, this particular position, the nucleotide did not match, whereas it matched everywhere at the remaining L-1 positions.

In that case, these two individuals are not isogenic. And mathematically, we define these two individuals as something that is called the Hamming distance. So, Hamming distance is defined

between two sequences. Both these sequences have to be of the same length. If they are of different sequences, then Hamming distance is not defined of the same length, in this case, L. So, HIJ is the Hamming distance between sequence I and J.

Between sequence I and sequence J. So, in the first case that we discussed, where the sequence was exactly the same at every one of these L positions, the Hamming distance was 0 if two sequences are exactly identical. And in the second scenario that we discussed, where the two positions were identical at L minus one positions but they differed from each other at this one position, then the Hamming distance is equal to one. So when we say that we have an isogenic population, What we mean is that if we take any two individuals in this population out of these N, HIJ between those individuals is zero. Their DNA is exactly the same.

For instance, if I have a DNA sequence ATGCCC, this is from individual I. And if I take individual J, its sequence is ATGCCC, and then obviously the Hamming distance between them is just zero. However, if I take individual K, its sequence is ATCCCC, then the Hamming distance of JK is equal to one because there is this one position out of six where the nucleotide does not match. And obviously, the Hamming distance between, I'm sorry, this shouldn't be IJ, this should be JK.

Hamming distance between JK. And obviously, if we were to do the same exercise for I and K, the Hamming distance IK would again be one because there is this one position in the sequence where the nucleotides are different from each other. So isogenic describes a scenario where the Hamming distance between any two individuals is equal to zero. So now let's look at this following experiment. Let's imagine that we have a flask.

And in this flask, I add 100 ml of media. of liquid media. The idea is that I'm preparing this liquid media so that the growth of bacteria can be studied, and I can perform some experiments associated with the growth of bacteria in this flask. And for that reason, this liquid media has to have all the ingredients that will facilitate the growth of E. coli under these given conditions. So typically, this would contain sources of

carbon, nitrogen, phosphorus, sulfur, oxygen, some trace elements, some amino acids, some vitamins, and so on and so forth. Water, obviously, hydrogen, and so on and so forth. And there are different ways of making these liquid media recipes to facilitate bacterial growth, but let's imagine I have 100 ml of that. To this flask, I add one bacterium. Now, experimentally, this doesn't really make sense.

It's really, really hard to start a culture by just putting in one bacterium. As when we study experiments associated with microbiology and evolution, we will see that we typically add millions

of bacteria to start a culture. But for the purpose of this discussion, let's imagine that we seeded this culture with only a single bacterium and we were somehow able to do that. Now, that is not enough. We need to give it 37 degrees Celsius.

That is the optimal temperature for the growth of E. coli. We will provide it with some sort of shaking so that this culture is maintained in oxygenated conditions. This just allows for better growth, and we will cover this later. To reduce the chances of any environmental contaminants going into our culture and competing with the organism that we want to study. That's another aspect of microbiology research that you have to be careful of: any contaminants coming into the culture because you really want to understand your organism under those precise conditions, and obviously, you have to make and maintain sterile conditions for that.

So under these conditions, if the nutrient conditions are just right, the temperature is right, and oxygenated conditions support growth vigorously. What's going to happen is that as this experiment proceeds, the bacteria are going to divide. It's going to divide first, and then we will have two progenies. And then these two will divide, and then we will have four. And this process continues.

And if we plot, let's say we started this experiment at 5 p.m. And then after seeding the culture with this one bacterium, we went home. And we are going to come back tomorrow morning at 9 a.m. Obviously, the culture is going to be just full of these bacterial cells in the flask. And in an experimental context, it's very easy to see because the liquid that we use for growth media is transparent.

It's a clear liquid. And when it's full of bacterial cells, it's going to be very murky and opaque; it can't be seen through. So let's say this is 5 p.m. When we seeded the culture. This is 9 a.m.

When we come the next morning. And this is the number of cells. So take a few seconds and try to sketch what this graph would look like, according to your intuition of how growth would take place. And just take a few seconds to go through this and try to sketch it for yourself before we continue this discussion. Okay, so what would actually happen in a scenario like this is something like this.

And this is referred to as the three phases of bacterial growth. The first one, where there is hardly any growth, is referred to as the lag phase. This one, the first, is called the lag phase. In the lag phase, there is very little actual cell division taking place. So imagine that the carbon source in this environment is lactose.

Now, this bacteria that I introduced doesn't have lacY and lacZ protein molecules to import lactose and then break it down into glucose and galactose. Hence, the first thing that the bacterium needs to do is make these proteins. Make them in sufficient quantity, import that lactose, start breaking it

down, start using it for energy and growth purposes, accumulate enough metabolites, make enough proteins, grow in size, and only then would this bacteria divide into two, and all of this is going to take some time. So this lag phase is associated with the time that is needed for the bacteria to acclimatize to the new environment that it's been asked to grow in and make the necessary cellular machinery which would facilitate that growth, and then you would begin to see growth. So that's the lag phase associated with it.

The second phase is called the log phase or the exponential phase. In this phase, the bacteria have accumulated the necessary protein molecules required for growth in that environment, and growth is taking place exponentially because you started with one, this one got And now this is ready to divide. It divided.

You have two. These will give four. Four will give eight. Eight will give 16. So if you plot this against time, it's going to be an exponential increase.

And that is what this phase of the graph captures, that this is an exponentially increasing number of individuals in the population. So this is the exponential increase time. In the number of individuals. However, what happens is that this exponential growth cannot be sustained forever because this is an environment in which we are conducting this experiment. This has finite resources.

With these finite resources, when you hit When you hit the middle of the exponential phase, what is going to happen is that there are far too many cells now. They are all competing for the remaining resources. Resources are decaying because as the number of individuals increases, the carbon biomass increases, and all the carbon and other nutrients have been taken from the media to make these organisms. As a result, the carbon available for the individuals present in the culture now

That number is increasingly small. Hence, the growth rate begins to slow down. So there is this slowdown phase at the end of the exponential phase. And eventually, you reach this stage, which is called the third phase, which is the stationary phase. And the stationary phase is something that is a bit of an enigma.

It's not as if the overall number remains constant, but this is because the number of births taking place is matched by the number of deaths occurring in the culture media. And as a result, this number stays constant, but there are really no nutrients left in the flask anymore. So this is how these numbers would grow and eventually reach something called, let's call it N final or NF. Now, 100 milliliters of culture media in this flask cannot hold any more individuals than NF. And for that reason, NF is called the carrying capacity factor.

of an environment. Typically, in literature, the variable that you would see used for carrying capacity is K. So, this is K or NF as we described it. which just means that in these environmental conditions, with this particular species of E. coli, if you were to grow this at 37 degrees Celsius, it cannot support more than K number of individuals. And that is what the population moves towards. So in this growth curve that we have,

Let's ask a couple of questions. We started with the number of individuals as one. At some point, that number reached n naught. We want to answer two questions associated with that. The two questions are, one, how many cell divisions took place for the population to transition from

From population size n equal to 1 to n equal to n naught. That is the first question. So, I will give you again a few seconds to think about it, and I will tell you the answer, but let us take 15 seconds. Think about how many cell divisions took place for this transition to take place. Okay, so I hope some of you got this.

But the way to think about this is that when you have a cell and it divides, it gives rise to two cells. And in that process, when we have these two cells, this one is no longer there. I have these two cells. So what happens that pre-division, if I'm looking at just this event, Pre-division, the number of individuals in the population was one, which was this cell.

Post-division, the number of individuals in the population was two. This is one cell division event. What you should realize at this point is that one cell division increases the population size by exactly one. Increases the population size by exactly one as a result of that population increased by one so For the population to go from n equal to 1 to n equal to n naught, how many divisions need to take place?

Every cell division would increase the population size by one. So, one cell division would increase the population size to two. Then, as these two divide, that means two more cell divisions take place. So, after three cell divisions, the population size would go to four. We can also think of it like this: we started with this one individual.

We have these two individuals after one cell division. Then this one would divide. That's two cell divisions, and the population size is three. This, this, and this. But then this would also divide.

So now the population size is three, but three cell divisions have taken place. This cell divided first, and then the two progeny divided, leading to three cell divisions and a population size of four. And now this will divide. So that's the fourth cell division, leading to a population size of five. And so on and so forth.

The point you should see, the pattern that you should see, is that one cell division adds only one cell. Hence, to go from n equal to 1 to n equal to n naught, I am increasing the population size by n naught minus 1. And since every cell division increases the population size by only one, I need to have N naught minus one number of cell divisions for this transition to take place. And I hope that makes sense to everybody. The second question that we want to ask associated with this is that as this change happened, just to go back to the previous slide, so we can plug some numbers.

That means to go from one cell to 100 cells, we needed 99 cell divisions. To go from one cell. To a thousand cells. We needed. 999 cell divisions.

And so on and so forth. N naught minus 1. And that is all we are doing here. To answer the. Second question.

The second question is as follows. That as the population Transitions from N naught. N equal to 1.

To N naught minus. n equal to n naught, as this transition takes place, how many generations passed? How many generations did it take for this change to take place? And to answer this question, we should look at it in the following way. That we start with one individual and let us call this generation 0 at t equal to 0.

After one generation, this individual would divide into And the population size at generation one, so let's say this is the number of generations and this is population size. Let me rewrite this. So in this representation of the system, the variables we are tracking are what is the generation number And what is the population size?

These are the two variables we are tracking as the system moves forward. At this point, the generation number was 0, equal to 0, and the population size was 1 because this was the only individual we had. After one generation, we get 1 as the number of generations and 2 as the population size. After One more generation, both these individuals would have divided, and the population size would be four.

After two generations, the population size is four. And let's just do one more, and then we'll try and see what is the pattern that's emerging here. So, after the third generation, we get Four, five, six, seven, eight. After three generations, we get eight individuals.

And as you can see, what is happening is that with every generation, the number of individuals in the population is simply doubling. So if we started with one individual, so in generation zero, we had one individual, In generation 1, we will have 2 times that individual. 2 times the population in

generation 0 because with every generation, the number is doubling. So, we get 2 into what was the size of the population in generation 0?

That was 1. 2 into 1, that is 2. At generation 2, the size of the population will be Two times the number of individuals in population one in generation, not population one in generation. Generation 1, and in generation 1, the number of individuals was 2.

So, this is 2, and we get 2 into 2, which is 2 squared or 4. Similarly, at generation 3, the number of individuals would be 2 times the number of individuals in generation 2. This number was 4, or let us say 2 to the power of 2. So, 2 into 2 to the power of 2. is 2 to the power of 3 or simply 8.

And we can keep on building this logic until we realize that after k generations, the number of individuals in the population will be 2 to the power of k. And we can verify this by checking here that after one generation, the number of individuals was two to the power of one. After two generations, it was two to the power of two. After three, it was two to the power of three. And after k, it will be two to the power of k. So that's what we understand from here.

So, how do we answer the question that we started with? That starting population... was 1. After k generations, the population size was 2 to the power of k. And I want to find out how many generations it took for the population size to become n naught. Which means I want a value of k such that 2 to the power of k is equal to n naught.

And that is the equation that I have to solve for the value of k. And if we just do this, we can take the log of both sides, log to the base 2, and we get k equal to ln of n naught base 2. And that is it. So, suppose we have, suppose we were asking how many generations it takes to go from 1 to 100, right? The answer is simply ln base 2 of 100, which, if you work out, is 6.67.

Another way to think about this is after one generation, this one individual will divide, and the population size will be two. So that's one generation. After another generation, the population size will be four. After three, it will be eight. After four, these eight divide.

So it will be 16. After five, These 16 divide, so it will be 32. After 6, it will be these 32 divide, so it will be 64. And after 7 generations, it will be these 64 dividing to 128.

But my question was, how long does it take for the population to reach a size of 100? And now, if we trace these generations, this is generation one. After two generations, it was four. After three, it was eight. After four, it was 16.

After five, 32. After six, 64. So clearly, six generations is not sufficient for this population size to reach 100. But seven generations is too much because, after seven generations, the population size reaches 128. And our question is only concerned with the population size reaching 100.

That means the answer has to be between 6 and 7. So this answer, 6.67, we can arrive at that by just doing this sort of simple calculation and realizing that the answer lies somewhere between 6 to 7 generations, which is the time it will take for the population to reach size 100. And so on and so forth. We can, I leave it to you as an exercise, but we can check how long it will take for this population to reach a thousand individuals. And that number of generations should be very close to ten.

For a 100 fold increase in population, the number of generations can be found by finding n in,

$$2^n = 100$$

Or,

$$n \sim 6.67$$

We will continue this discussion and try to get an estimate of what the diversity of genotypes is that accumulates while these generations and cell divisions are going on when we continue this in the next video. Thank you.