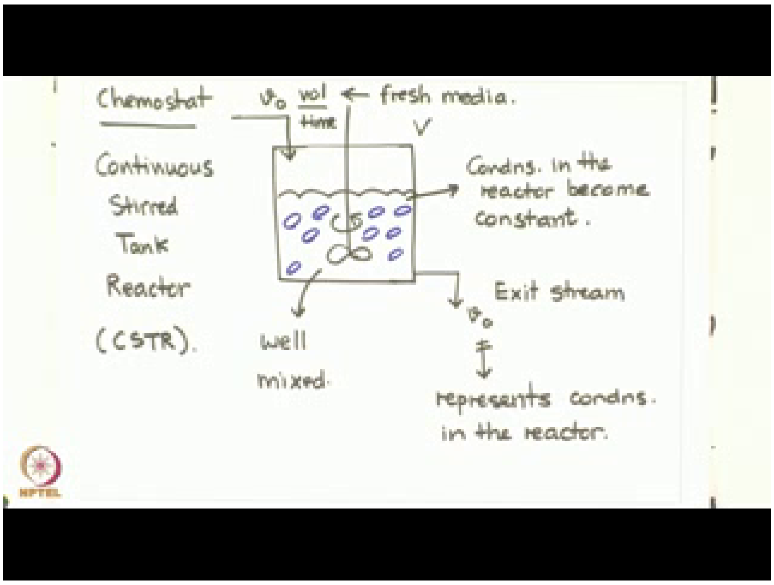
Introduction to Evolutionary Dynamics
Prof. Supreet Saini
Department of Chemical Engineering
Indian Institute of Technology, Bombay

Lecture – 28
Estimating timescales of Evolution

Hi everyone, continuing our discussion from the previous time where we were discussing one of the ways you can do a evolutionary experiments in lab with bacteria. We will start our discussion with the second way you can do these experiments in a lab and then move our discussion forwards as to what is the kind of information that we can look at from these experiments and theoretically try to come up with analysis which might tell us something about these experiments and making us being able to predict what might happen in these populations.

(Refer Slide Time: 00:58)



So, the second one the second we you do these experiments is using something which is called a Chemostat. So, Chemostat is as is essentially just stored reactor. So, you have what this means here that this the contents of this reactor are well mixed which means there are no special heterogeneities associated with concentrations, concentration of every specie nutrient inside this reactor are constant everywhere and you have bacteria growing in this environment.

Now, the problem with the test tube experiment is that as we had discussed last time that you grow test tubes at a particular time you add n naught number of individuals and you let growth happen for a day, and next morning you come to the lab and you transfer and the number of individuals in the test tube has gone up to n t and from these n t number of individuals you pick up n naught and transfer them to the next tube and the process repeat cells self every single day. The main problem associated with this type of experiment is that the condition in which growth happened for those 24 hours does not remain constant. In the beginning few hours of the experiment the conditions of growth for bacteria were very good they were not very high in number, the nutrients were plenty and hence growth happened at a particular dynamics.

However as you move forward in time for that one day in which growth takes place in a test tube the conditions associated with the experiment change quite dramatically nutrients become depleted the waste from metabolism of the individuals gets recreated in to the environment, you have cells entering stationary phase of growth and as a consequence of all of these you have physiology associated with the cells changing with time. That is not really ideal for an evolutionary experiment because one of the main things that you want to test is that how do my bacteria adapt when I grow them in this completely defined environment e that I am making in this test tube.

However because of the points we just mentioned that environment e does not remain constant in a test tube type of a setting. Hence we move to something which is called a continuously stirred reactor or a Chemostat which elevates that problem and helps you maintain constant environmental conditions at all times which truly gives you a measure of how will bacteria adapt in this defined environment.

So, this reactor is going to have a volume V associated with it and it has a particular defined environment and you have stirring taking place to make sure that there are no heterogeneities of concentration in space, but this still looks like a test tube experiment that we talked about the only difference being that there is a stirrer and this is larger in volume, but what elevates that problem is that you have an incoming stream to this reactor at the rate v naught volume per time which is carrying fresh media.

So this is the media which you desire to do your experiment in and you have a continuously you have this stream carrying these nutrients continuously at rate v naught.

So, this would mean that the volume of this reactor increases at the rate v naught, but we elevate that you have an exit stream which is removing media and cells from this reactor at exactly the same rate v naught. If you do this for a long enough time what these conditions ensure is that the conditions inside this reactor. So, when you do this experiment for a long time analysis of a reactor such as this shows that the conditions in the reactor become constant. They do not change with time and by conditions we mean the associated concentration of every nutrient what is the number of cells in the environment and so on and so forth, every single one of those aspects associated with life in this reactor becomes a constant with time.

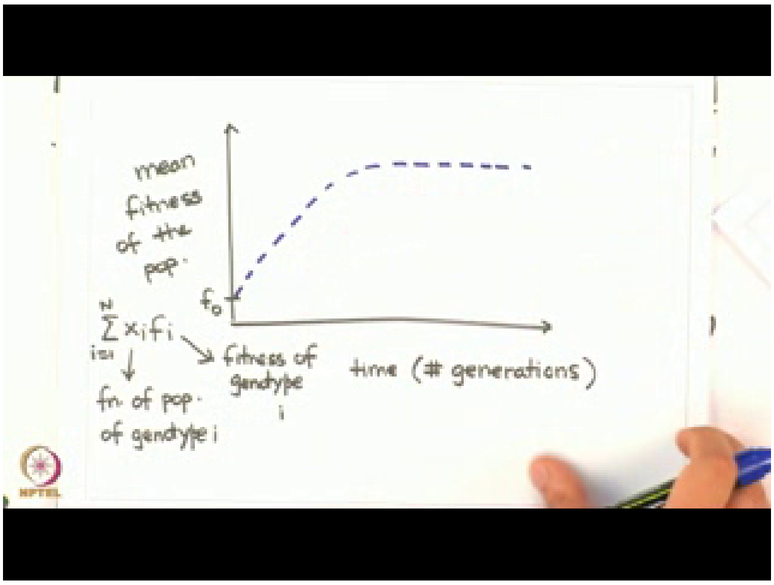
And what is more the conditions are equal to that in the exit stream. So, this represents conditions in the environment in the reactor because the analysis of this reactor would show that the exit stream is carrying what is present well mixed inside this reactor. So, now, you have this reactor like setting which gives you constant conditions with time because of these exit and entry streams one of the consequences you should realize of this is that the fresh stream that is coming in that is only carrying the incoming stream that is coming in to the reactor is only carrying fresh media it is not carrying any bacteria.

Whereas the exit stream which is again going at going out from the reactor at rate v naught is carrying bacteria and spent media from inside the reactor. And what we are saying is that if we do this experiment for a long enough time the reactor establishes a steady state associated with it. What that means is that the number of bacteria which are removed by the exit stream which is flowing at volume v naught per time rate the number of bacteria removed by this stream per time has to be exactly matched by the number of bacterial divisions taking place inside the reactor such that the number of bacteria inside the reactor does not change with time. So, that equilibrium also gets established between these two variables. So, you also know the rate at which division is taking place inside the reactor at any given point in time.

So, again you do these experiment for a long enough time you maintain your incoming stream for a long enough time and you keep taking periodic samples of what is present inside the reactor from that exist stream and that keeps you a snapshot into what is happening after 50 or a 100 generations or after; however, many generations you may choose to take your freezer stocks. That is a second way that you would do your experiment. This is also called, the analysis of this type of a reactor is also referred to as a continuous stirred tank reactor which is CSTR.

Of course you are doing an experiment like this we would mean that you have to maintain sterile conditions for the reactor for; however, long that you plan on doing the experiments, but it affords you this great advantage that the environment associated with the experiment is now truly constant which was not the case with the test tube setting that we discussed in the previous lecture.

(Refer Slide Time: 09:25)



So, now, if we do an evolutionary experiment via any of the two ways that we have discussed and try to answer a simple question that in an evolutionary experiment what do I expect will happen to this graph on the x axis I am plotting time and let us again represent that n number of generations and y axis we represent fitness, mean fitness of the population. Which means if I have n different genotypes the mean fitness is simply represented by sigma x i f i where x i is a fraction of population of genotype i and f I is the fitness of genotype i. This is true if you are if a population is carrying n distinct genotypes in the environment.

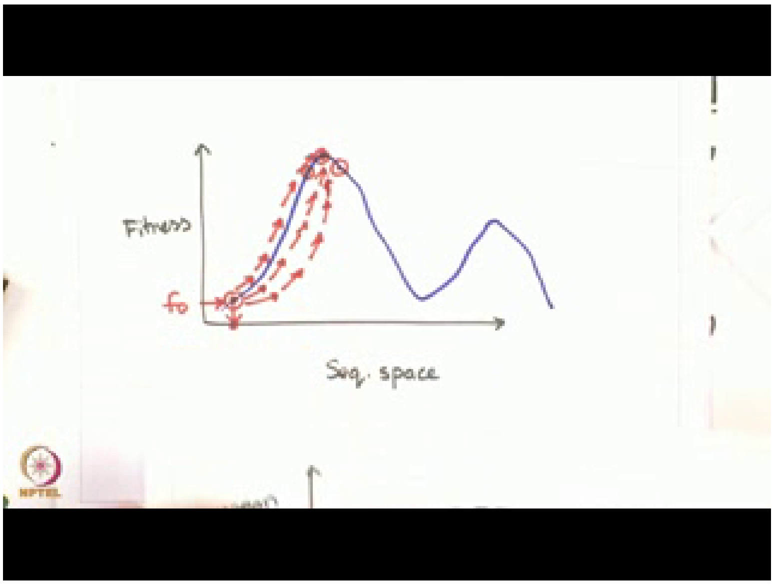
So, we start our experiment by set of individuals which are isogenic and are all at fitness f naught. So, if you are starting with a genetically identical population where all individuals have this fitness f naught then the mean fitness of the population at that time is also going to be equal to f naught. Starting from this point I would like you to sort of pause this video for a minute or so think about this graph and qualitatively sketch as to what you would expect as you were forward in time to happen to the fitness variable associated with the population.

So, just may be take a two minute pause think about this and try to come up with intuitive solutions to this graph that you can think of. You can think of more than one ways you can think of more than one logical ways or logical graphs that appear reasonable to you and try and come up with reasons that you would expect to happen to this graph.

So, from a number of experiments evolutionary experiments that people have done it is found that the nature of this graph as we forward in time looks something like this. What is going to happen is that initially you have a lot of beneficial mutations which are available to you. So, the individuals keep picking those beneficial mutations and the mean fitness keeps increasing with time.

Eventually however, you run out of beneficial mutations to be conferred to the population and then the fitness can no longer increase and it platoons with time. And now we were to think about this is in the context of fitness landscapes that we have discussed earlier in the course. So, remember during fitness, when we are defining fitness landscape you have sequence spaces which are connected by edges and nodes - nodes represents sequences two sequences are connected if they differ only by a single nucleotide and the third axes corresponds to the fitness associated with that particular genotype represented by a node.

(Refer Slide Time: 13:04)

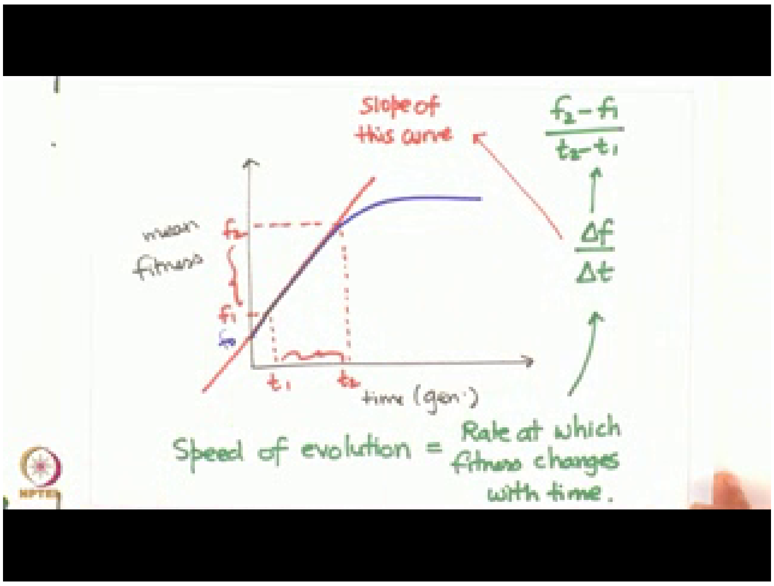


So, in a simplified one dimensional sequence space this is fitness we had seen that we could get many types of fitness landscapes, but let us imagine that this is the fitness landscape that I am talking about. And now in the evolutionary experiment that I am starting with this is the point where fitness of the n individual was f naught and this is the genotype of my starting cells. Everybody in the starting population had this particular genotype and this particular fitness.

Now, what is going to happen is that individuals are going to keep acquiring these beneficial mutations and the mean fitness associated with the population keeps increases and eventually as we saw because of finite mutation rate there will be a quasi specie equilibrium between the mean fitness between the peak fitness and the neighboring sequences which are very close to the peak fitness. So, that quasi specie equilibrium will establish itself very close to the peak of the fitness landscape and that is what is happening in this intuitive picture that we have drawn here. When a experiment is just starting there are lot of beneficial mutations which are available to this help. So, instead of going this way the population could also move along this phase of the curve or could also move along this phase of the mountain and so on and so forth.

So, these mutation, there are lots of mutations available which increase mutation which increase fitness and the cell is continuously able to access those mutations and the rate of the fitness is sort of linear with time as we move forward in generations. However, as we approach the peak the number of mutations which further increase fitness of the population drive up and that leads to this plattoying effect of in fitness and fitness does not increase with time as it was during the initial phase of the experiment. So, you have this saturating kind of behavior that is being that is observed experimentally when you monitor fitness with time in a defined environment.

(Refer Slide Time: 15:40)

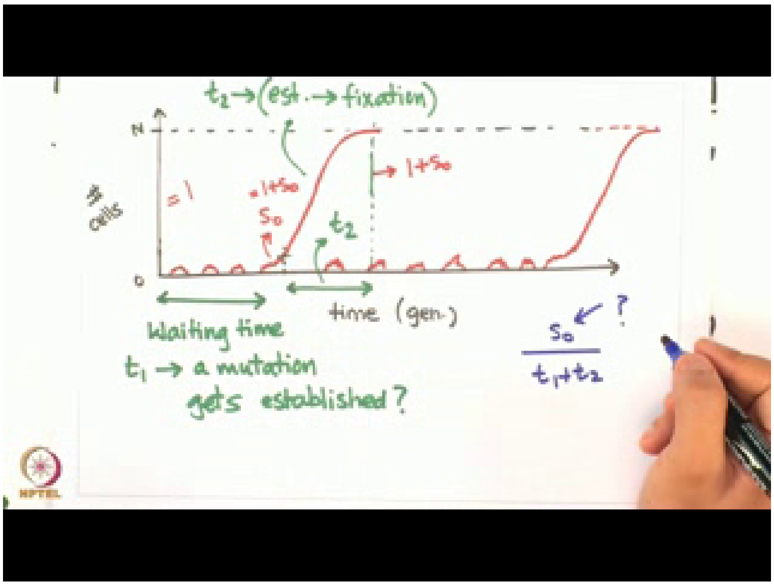


But what we can see here is that in these representations we can define a variable called speed of evolution and let me just redraw that graph again. So, this is time in generations and this is mean fitness of the population and we note that starting from f naught the population increases and then saturates. And one of the ways we can define a quantity called speed of evolution is equal to rate at which fitness change changes with time. If fitness is changing very rapidly with time; that means, you have rapid evolution taking place because the population is moving towards higher fitness is in a very very fast way; however, if change in fitness with a change in time is happening very slowly; that means, evolution is not taking place at such a phase. So, this can we defined as delta f divided by delta t where delta f is just f 2 minus f 1 fitness at time 2 minus fitness at time 1 divided by t 2 minus t 1.

And what this tells me is that if I were to choose t 1 and t 2 and on this graph find out its corresponding f 1 and f 2 then this change f 2 minus f 1 divided by this change t 2 minus t 1 that is what gives me an estimate of speed of evolution associated with the population. And this can be seen that this quantity is just equal to slope of this curve. Another thing you should note here is that in this region of the curve the curve although it becomes non-linear and has this saturating tendency in this region of the graph the curve can be estimated as almost a straight line. So, in this straight line region we say that evolution is happening linearly with time as time increases by a certain rate fitness is increasing by a rate which is proportional to the increase in time and later on as you move towards larger times that rate slows down and eventually comes to 0 when this curve perfectly becomes horizontal.

But we will be talking of this linear region where evolution is taking place at a linear rate or fitness is increasing at a linear rate with time. So, let us try and develop expressions associated with that how do we even go about thinking something like that. So, in an experiment if I am doing an evolutionary experiment, let us go back to one of our representations and see what does that tell us.

(Refer Slide Time: 19:06)



So, we will go to the first representation that we did where x axis represented time in generations and y axis represents number of cells from 0 to n.

So, first thing that I want to imagine is that how would evolution happen in this chemostat type of a setting where I am doing my evolutionary experiment. If I want to just imagine this in my head I know that there are lot of mutations are going to get lost and those are going to represented, those are going to be represented in my graphical setting as follows. I have a beneficial mutation happen and it goes away by drift this now represents in my chemostat as those mutations being washed away from washed away from the reactor to the outside before they had a chance to spread into the population. These are beneficial mutations which are happening, but were not able to spread because of drift.

After some time I have another beneficial mutation not surviving another beneficial not surviving and so on and so forth eventually I will get one beneficial mutation which is able to get established you should always remember the three fates of a population extension, establishment and fixation. So, eventually you have a mutation that gets established and once it is established; that means, random chances are not going to be able to eliminate that mutation and now selection takes over and selection ensures that this mutation which has gotten established which by definition is beneficial in nature is able to go to fixation. So, you have this establishment taking place and then you have mutation going to fixation.

And then the fate repeats itself that you keep having these beneficial mutations, but they do not survive and eventually after while you have another beneficial mutation which may survive and get established and once established it goes to fixation. So, that is the intuitive picture that we have of evolution taking place in a of evolution taking place of a microbial population in experiment. Now what we want from this type of an output is the slope of this curve which gives me the rate of evolution or the speed of evolution associated with this microbial population that is what going on to work towards.

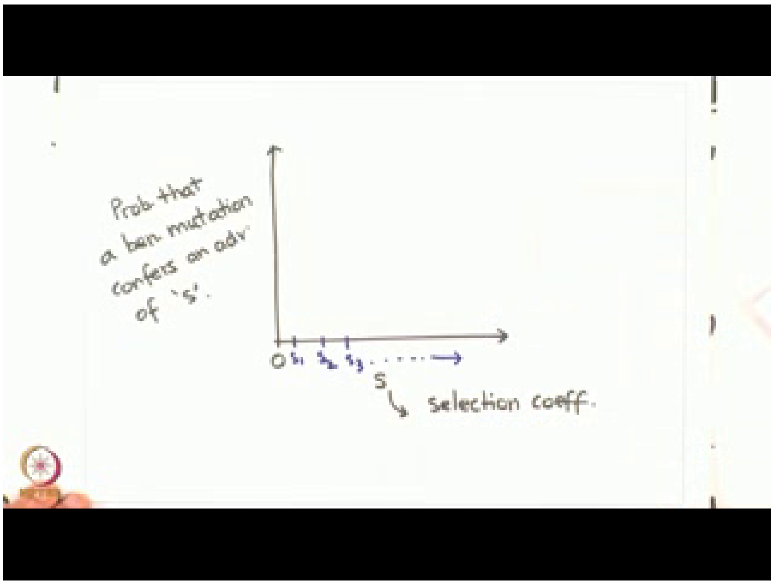
So, the first thing that we want to understand is there are two quantities here that what is the waiting time, what is the waiting time t 1 such that a mutation gets established. Remember a mutation getting established is referred to as the number of individuals which belong to that particular genotype exceeding 1 by S where S is the selection coefficient associated with that particular mutation. So, what is the time that is associated with mutation getting established? We are interested in this because if a mutation is established such as this one it means that it can no longer be eliminated because of drift and now selection ensures that the numbers associated with its frequency got to one and the parent genotype is eliminated. How long do I have to wait for this to happen?

The second time that I am interested in is let us call that t 2 and that time is what is the time it takes for the mutation to go from establishment to fixation. How much time does a mutation need to go from this establishment event take place here and the fixation event take place here this time is referred to as t 2. So, these are the two times that I am interested in, which will tell me because I can see that fitness of the fitness of the population has gone up by the magnitude of this beneficial mutation fitness of all the individuals has gone up. So, let me rephrase that. At the beginning of this experiment at this time the fitness of every individual in the population was 1.

Suppose the mutation that happened here has a selection coefficient S naught associated with it; that means, fitness of every individual which is carrying this mutation is equal to 1 plus S naught if that is the case when the mutation gets fixed in the population at this point fitness of every individual in the population now is 1 plus S naught. Since the fitness of every individual has changed from 1 to 1 plus S naught; that means, the mean increase in the mean increase in the fitness of the population has gone up to one plus S naught and the mean has increased from 1 to 1 plus S naught; that means, a change of S naught has taken place in time t 1 plus t 2 and that gives me the speed of evolution in that sense. So, I need to understand that I need to get an estimate of t 1 and t 2 and I know that in the time t 1 plus t 2 S naught fitness changes by S naught, so in that sense speed of evolution is just given by S naught divided by t 1 plus t 2.

But before I get that I have to understand that how our mutations happening in the population what is the probability associated with mutation of which confers a fitness of S naught which confers a fitness advantage of S naught to an individual. Intuitively what do we believe?

(Refer Slide Time: 25:52)



Let us draw in pictures this graph, on the x axis I have S which is the selection coefficient and this graph starts from a value of S equal to 0, I am not interested in negative values of x because negative values of x mean that there is a deleterious mutation that has happened in the environment in this mutant. And deleterious mutations are not going to get selected for the selection will make sure that they are eliminated hence I am not going to worry about deleterious mutation I am only interested in beneficial mutations. Beneficial mutations which would mean S is positive.

On the y axis what I want to represent is that probability that a beneficial mutation confers an advantage of S. So, what we mean here is that this is a probability distribution curve and what we mean here is that this is S 1 this is S 2 this is S 3 and so on and so forth this is increasing S here do we un dissipate that every time a beneficial mutation happens its is its magnitude more likely to be S 1 or S 2 or S 3 or is its magnitude likely to be equally distributed or are beneficial mutations which confer a large advantage in fitness are more likely to those which confer a small advantage and so on and so forth. The essential question is what is the frequency distribution associated with beneficial mutations that we know of.

And this is important because when we are interested in rate of evolution we want to understand this graph and this graph tells us that the speed of evolution can just be given by S naught divided by t 1 plus t 2. But we do not we do not really understand what is S naught more importantly the probability that any mutation is able to survive drift and is going to establish itself this probability is itself dependent on S and in fact, that probability is equal to S have you saw a couple of lectures back..

So, understanding the fitness advantage conferred by a beneficial mutation is extremely important for us to be able to comment on the nature of the dynamics associated with a particular mutation. Before we end this lecture I want you to just imagine this scenario that if you were to think of the magnitude of a beneficial mutation on x axis such as shown on this graph and on the y axis the probability distribution of a beneficial mutation of a particular magnitude happening what do you imagine that graph would look like. And this graph is understanding this graph is important because of very obvious reasons that we just discussed.

So, again, after end of this lecture I want you to think about this graph and draw intuitive pictures that that agrees with the understanding of an evolutionary process that you have in your mind and we will start the next lecture starting with this, starting with the discussion on this particular graph.

Thank you.