Evolutionary Dynamics

Supreet Saini

Chemical Engineering

Indian Institute of Technology Bombay

Week 12

Lecture 60

Thank you. Hi, everyone. Welcome to the last video of the course. We'll discuss a couple more interesting studies, I think, about how evolution experiments are done in the lab and what sort of light they shed. And then I will discuss a couple of broad problems that we haven't had a chance to discuss in this short course.

So, last time we discussed how synonymous mutations alter—just to briefly summarize what we did—that synonymous mutations can have fitness effects. This is not to say that every synonymous mutation will have a fitness effect. Synonymous mutations can have fitness effects. And they do this by changing protein amount. This protein amount can be changed in several different ways.

First, by altered stability of mRNA. As a result, if stability decreases, that means you have less mRNA and you get less protein. Secondly, you can get that by altered access to ribosomes. If at the five-prime end of this mRNA, we have a tight hairpin or the ribosome binding site is embedded deep in the mRNA structure, then it's harder for the ribosome to get access. And lastly, the use of a rare codon slows down translation.

Secondly, and somewhat surprisingly, synonymous mutation can also change protein structure. And this we saw was because of an altered co-translational folding kinetics. And for both these studies, we know of several examples which illustrate this point. And again, I want to emphasize that is not to say that every synonymous has a fitness effect, but many synonymous mutations do have a fitness effect. In a study that was done a few years ago, it illustrated this point very nicely.

what they did was you have DNA and then transcription takes place and you get mRNA. And this mRNA is synthetically created in such a way that the first three-fourth encodes for GFP, that means the green fluorescent protein, but not the full GFP, let us say only

three-fourths of the GFP. Then you have one-fourth You have the one fourth remaining part. And after that, you have RFP, which is red fluorescent protein, but not the full RFP, but three fourths.

And now what the authors did that in the first scenario, you place rare codons here. After the three fourths GFP, you place rare codons here. And what these rare codons do is that when ribosome comes here, there is a long pause because it has to find the appropriate tRNAs which are carrying these amino acids which are being encoded by these rare codons. As a result, the GFP amino acids, the three fourths GFP amino acids that have been encoded here, they will not stay like this. They will fold and we get a protein structure which is encoded by three fourths of GFP.

And this doesn't fluoresce. So there is this is GFP, but not full GFP and it doesn't fluoresce. Eventually, these rare codons are being translated and then the rest of the transcript is translated first and this remaining bit when it is translated. So, now we have this GFP, three-fourths GFP and the rest of the codons. which are encoded by this one fourth of GFP and three fourths of RFP.

Now, interestingly, this one fourth is such that it will fit in with GFP and make the GFP complete or this one fourth can also fit in with RFP and make the RFP complete. So, in this particular design, when rare codons are here, GFP folds and this one fourth is then unable to incorporate into this already folded GFP structure. So what happens now is that this rest of the amino acid chain folds and we get the following structure that we have a three fourths GFP and a fully folded RFP. Excuse me.

And three fourth GFP. So because because this cannot fluoresce. Because this cannot fluoresce when I look at the cells. When I look at the cells, the cells are fluorescing red. So, when I look at these populations under a microscope, the cells will appear as if they have red fluorescent protein, which is being made because of these following reasons.

However, The authors of this study used this altered design where now we have this DNA and the resulting mRNA. This is the part which is three-fourths GFP. Then we have one-fourth GFP and then we have three-fourths RFP. Now, instead of placing the rare codons here, the authors placed the rare codons here.

This means that when this transcript is going to be translated, the halt, the ribosome halt, is going to occur when it's here. And the first part of the transcript is going to be translated reasonably fast. As a result, the amino acid chain is going to be now that the

translation process has halted. This amino acid chain will now fold because there is a long pause when the ribosome is not moving. But this one-fourth GFP will now work with the three-fourths GFP and form the fully folded GFP.

So, this is full GFP because this one-fourth is now working with this three-fourths GFP. And once these rare codons, these are the rare codons. Once these rare codons are translated, then translation resumes and the rest of the protein is then translated. And when the ribosome comes here, this is the rest of the protein. And eventually this will translate and the net result of the folded structure would be this, where this is three-fourths RFP.

And this, of course, will fluoresce green. And all of this, interestingly, is accomplished by only changing the location of these rare codons. In the previous case, you placed some rare codons here and got a different protein shape, making the cells fluoresce red. In this case, you changed the position of the rare codons to after the one-fourth part, making the cells fluoresce green. So these synonymous mutations can not only change protein amounts via these different mechanisms we discussed but also protein shape.

And protein shape ultimately determines protein function. Hence, a synonymous mutation can also change protein function. This, I think, is an extremely interesting result that has emerged in the last few years. All right. Let's quickly...

Also, discuss one example of what happens in a mutation accumulation experiment. That is the experimental technique we did not quite discuss after explaining how it is performed, but what observations emerge from a mutation accumulation experiment. So, let us imagine we have E. coli. The way we will conduct these experiments is by growing them in a tube.

Then I will take a very small volume, V naught, and this V naught, let's say, only contains about 100 individuals. I will spread these 100 individuals on a plate. Of course, I can't see individual E. coli, but when I spread them, my calculation shows that this much volume of the liquid must have 100 individuals. So when I spread them, I know there are 100 individuals there, but I don't know where they are. So they will be spread somewhere whose location I don't know.

So there are these 100 individuals spread over a petri plate. What I do is I want to pick one of them. But I don't want to make any biased choices. So when I wait, I let this grow overnight in an incubator. Grow overnight at 37.

And what will happen is that for each one of these individual cells that was placed on a petri plate, It has access to resources, and the temperature is right. So, doublings will start to happen: one individual will become 2, 2 will become 4, 8, 16, and so on and so forth. By the time I give it 14 to 16 hours, these would be many, many cells, and they would look like little mounds like this on a plate, and each of these mounds is called a colony. Now, obviously, in this process, mutations have happened.

And I want to pick one individual from one colony completely randomly. Now, how do I randomize this process? Remember, mutation accumulation experiments are where I want selection to be absent. So everything has to be completely randomized; only drift has to decide the fate of this evolution experiment. So how do I do this?

So the typical technique is that I don't wait until this moment to pick an individual colony. That is because by looking at the colony, I will be biased. Some people like to pick a very average-looking colony. Some people like to pick big colonies, some small, and so on and so forth. Some colonies might appear different in color, shape, and so on and so forth.

This is all happening because of the phenotypic state of the cells in that population and the mutations that have been acquired. So you are attracted to a certain type of colony. And in order to sort of get rid of that bias, what we do at the time of plating is that I will mark a section of the plate completely randomly. So I don't know where these 100 cells are. But at the time of plating, I will mark—I will physically mark a part of the plate like I've shown here.

And the idea is that I am saying whichever colony comes inside this box or closest to the box is the one that I'm going to pick. That might be the biggest, smallest, most colorful, irregularly shaped, or most regular-looking colony. I'm not going to care about any of that. But whichever colony comes here is the one that is going to get propagated in the experiment. Everybody else is going to get discarded from the experiment.

So, in this way, the choice of picking this colony was randomized. Now, the choice of colonies has been randomized, but now I want to pick one individual from this colony completely randomly. I don't want to pick the cells at the periphery. I don't want to pick the cells near the center of the colony or at the top of the colony.

I want every cell present in that colony to have an equal chance of being picked and propagated forward. So, what I do to ensure that is I pick this colony up physically from

the plate and I And I put it in a culture tube and vortex it so that all members of the colony are spread in this tube here. Now I take V_0 amount from this. So when I take V_0 amount, my goal is to take about 100 individuals from this colony.

And now because every individual has been picked up from the plate, and suspended in this tube and I am out of that entire population, I am going to pick 100 individuals. Every individual has the same probability of being picked and every individual has the same probability of not being picked. So, this choice of which 100 individuals I picked from this colony in this way has also been made completely randomly. I did not just simply touch the periphery of the

of the colony and suspended. I tried to pick as much of the colony as I could. Of course, there will be some cells which will be kept there, but that is trivial as compared to what I picked. So, this has also been randomized and after picking this I do the same step again that I spread this on a plate and again I do not know where are these 100 individuals because I cannot see individual cells and at the time of plating I am also marking a plate. in the plate where I am saying that whichever colony comes inside or closest to this particular marked area is the one that is going to get propagated forward.

So, in this way, the choice of colony and the choice of 100 individuals that are being propagated in the population is completely randomized. Now, when I do this, what is going to happen is that this colony, every colony has 10 to the power 8 cells. That means lot of mutations have taken place here. And there will come a time that by chance these 100 individuals will also have individuals which have acquired mutations. And eventually, I will end up picking a colony which which which was founded by one of those individuals which was already carrying a mutation.

And so on and so forth. So as I move forward and do more and more plating, more and more mutations will get accumulated in this line of experiment. And hence, this is called mutation accumulation. I forgot the word accumulation, mutation accumulation experiment. we are accumulating mutations in a line, but without letting selection act.

So, what is a typical sort of readout from this experiment is that if we look at this, this process is called one transfer. And a typical mutation accumulation experiment is you will do this maybe 200 times, 300 times. And as we saw with Lenski's experiment, you will not do only one line, you will do several lines in parallel. Let us say you do 100 lines in parallel of like this and you are doing it, you are doing this parallel work for 200 transfers.

So, if that is the case, then what is the sort of a readout? One interesting readout is that I would like you to plot the following. Let's say this is number of transfers. And this is fitness. Let's say this is the starting fitness F naught, which is the ancestor fitness.

What I would like you to try and guess here is how the fitness would change as I go through this experiment for, let us say, 300 transfers. I will give you 10 seconds to think about it, and then we will discuss what is actually seen in a mutation accumulation experiment. So, let us normalize this ancestor's fitness to one. What is seen in all mutation accumulation experiments? Mutation accumulation experiments have been done on a variety of organisms, from C. elegans to E. coli to yeast to Drosophila, and every one of them shows this particular pattern.

That as populations are propagated, this is what happens to fitness. Fitness decreases almost linearly. And remember that we said we had eliminated selection from this experiment. So, if we had eliminated selection, why is fitness decreasing? It's decreasing because as this individual is being evolved in the population—let's say this was the ancestor genome—after going further ahead, it accumulated one mutation.

After going a little bit ahead, it accumulated another one, and another one, and so on and so forth. So, as it is moving forward, it is accumulating more and more mutations. Now, given a mutation, what is our null expectation? That given any random mutation, is it more likely to be beneficial or deleterious?

And I think we all recognize by intuition that deleterious mutations are more likely to occur as compared to beneficial mutations. So as this line is moving forward, it's more likely to acquire a deleterious mutation as compared to a beneficial mutation. I mean, some beneficial mutations will also get acquired, but by and large, more deleterious mutations will get accumulated. As a result, this fitness of this line, when I check it periodically, will be continuously decreasing. Now what I can do, so that's one, which is basically just a consequence of the fact that there are more deleterious mutations as compared to beneficial mutations.

What I can also do is that I can sequence these lines and I can try and find out what is the nature of mutations that are happening. How many are SNPs? So I will sequence this. I will sequence the ancestor that I started with and I will compare them and find out that what is the nature of mutations. How many are SNPs?

How many are insertion deletions? How many are duplications? And so on and so forth. So, I will quantify this and this gives me a very detailed information of not just about what is the mutation rate, but what is the mutation rate associated with each one of these types of mutations. So, that is another important piece of data that comes out of these mutation accumulation rates experiments.

And obviously, I do not do this in one line. We said that we are doing each line for maybe 300 transfers, but we are also doing 100 lines in parallel, just as Lenski did 12. So, each one of these lines is going to give me one trajectory. So, I do not have one line here. I actually have 100 lines here.

Mutation accumulation experiments conclusions are never drawn from one line because mutations occur randomly. And it's possible that by chance, one line could acquire more SNPs and may have no duplication. Whereas another line may have more indels and very few SNPs and duplications. But when we average these effects over 100 different lines, then we get a statistical measure of what is actually happening in the genome of the organism. Fine.

So that's the sort of input that we have, that we obtain from mutation accumulation experiments. And as I said, for a number of organisms, these types of experiments have been done for a large number of organisms. And the lessons that we draw from them are robust. The principles hold across organisms of different complexity. And a lot of what we know about mutation rates actually comes from studies such as this.

In the last few minutes of the course, I want to discuss a couple of aspects that we haven't had a chance to discuss. The first one is to do with the fact that ecological niches rarely have one species. So, ecological niches have much more than one species. In fact, there have been a lot of studies. So, what are the popular niches that people study? Isolating—so one of the popular ways of studying how bacterial species interact with each other is through studying bacteria that are present in soil.

And there have been studies that have tried to isolate different species from a very small grain of soil present in a garden, for instance. And what we find is that even in a very small grain of soil, the number of species present is in the several dozens. So, this particular sample from a study that came out about 10 years ago had 4 to 5 dozen different species—50-60 species coexisting in that grain of soil. Our group has isolated bacteria from soil, and from one gram of soil, we isolated 25 different bacterial species that were co-inhabiting that niche.

Remember, these are species that we have been able to isolate. There are many others that we have not been able to isolate. So, the number of species that are coexisting is very, very large. Now, this poses a few questions: Why is the fittest species not taking over? What are the principles?

That maintains this population structure. This could be thought of in the context of the soil microbiome or even the gut microbiome. Many studies take place these days, studying the microbiota present in our intestines and how species dynamics change. The second issue is stability—that These systems are exposed to perturbations all the time.

In soil, sometimes it floods; sometimes it's dry. Temperatures are high during the day. Temperatures drop at night, and so on and so forth. We drink a lot of water. We eat, but then we don't eat for hours, and so on and so forth.

So why? How is it that these systems These niches have dozens of interacting species, yet they maintain their numbers. They don't go extinct. How is the stability or robustness of these niches maintained?

So these are two important questions on which a lot of work is going on these days. The last thing that I want to discuss is the phenomenon called HGT, which stands for horizontal gene transfer. The idea of evolution that we have discussed in this course is that I, as an individual, divide into two progeny. And if division is faithful, then the progeny have identical DNA. But then somewhere down the line, one of the progeny will have a mutation.

So evolution in this sense is vertical. That evolutionary change takes place across generations. As you transition from one generation to another, evolutionary change can take place. A few decades ago, it was found that, in this vertical view of evolution, what is true is that individuals don't evolve; species evolve across generations. Because if I have a certain genotype, I am going to live with that genotype and not evolve.

But my progeny might have a different genotype, which might be better suited for the environment that we are living in. So I, as an individual, don't evolve in a biological sense at least, and species evolve through generations. That is the understanding that we have through this vertical prism of evolution. As against this, a few decades ago, it was found that bacteria have this phenomenon: if they are carrying pieces of DNA, they can trade DNA where this individual can actually acquire a little piece of DNA from that individual, and this trade of DNA can happen.

And the popular mechanisms of this are direct trade with another bacterium. Or viruses carry pieces of DNA around. So one virus might have infected another individual and is carrying its piece of DNA, and when it infects me, I inherit that piece of DNA. And third, from the environment. So many species, the most popular one studied is the soil bacterium called Bacillus subtilis, have this capacity: if in the environment there are pieces of DNA, it has the ability to pick up a piece of DNA and become this.

So through horizontal gene transfer, I don't need to divide, and I'm able to evolve as it is. So, this is horizontal because I can trade genes with my peers of my own generation and evolve in a biological sense. And this has huge evolutionary implications and a very interesting and dynamic mode of evolution because remember, here the kind of changes that are happening are small. There was a G here, there was a T here. So in most cases, this leads to small incremental changes.

There was a protein functioning; you made it better. But in this particular case, you can acquire completely new traits because you are acquiring large pieces of DNA from the environment through any of these three mechanisms and evolve. Again, lots of work is being done to understand this. Lots of experiments—interesting, cool experiments—are being done to understand this. But for the purpose of this course, we stop here.

I hope you've enjoyed these 30 hours. And I hope the purpose of these 30 hours is to just scratch the surface of this beast of a subject that we call evolution. And for those of you who are interested, I hope you will take up further courses to enhance your knowledge in this area. Thanks a lot. And we close the course with this.

.