An Introduction to Evolutionary Biology

Prof. Sutirth Dey

Biology Department, Population Biology Lab

Indian Institute of Science Education and Research (IISER) Pune

Week 5 Lecture 29

Migration: Numerical simulations and analysis

Hi, so in our last discussion we mentioned that most of the evolutionary forces we talked about were selection and drift. assortative mating, inbreeding all of them will end of the day tend to reduce the amount of variation that is present in the population. Whereas the one force that adds to variation in the population is mutation, It seems to be a weak force primarily because the rate of mutations, as we know, is not very high. So, there has to be something else, some other way by which organisms or populations will end up getting their variations. And as you might have guessed by now, that evolutionary force happens to be migration. However, before we go into the details of migration, we need to sort out a few definitional issues.

And the first thing is related to the term itself. Now, when we talk about the term migration in the field of ecological evolution, There are two slightly different ways in which we think about it. So, if you talk to a behavioral ecologist, they are going to simply say, "Look," Migration refers to any movement of organisms or their propagules, meaning seeds and other juvenile phases, in space. Now, this is also the way in which common people in their regular day-to-day usage will think about migration.

However, if you talk to an evolutionary biologist, they are going to think about migration slightly differently. They will define it as the movement of alleles or genes between populations. So, any movement of organisms that does not lead to a change in the genetic composition of the population, That is not migration according to evolutionary biology. So, just to give you one example to make this distinction clear, most of you are probably aware of this particular fish. This is the famous Chinook salmon.

So, as you know, these fish typically tend to breed in streams in the US, Canada, and other countries. So, they spend most of their lives in this ocean, but just before breeding, they come to these streams. They travel, you know, hundreds of kilometers upstream, and then they spawn and die. The juveniles hatch, and after hatching, they take a massive long journey all the way back to the ocean. They grow up over there, but again when the time comes to breed, they often come back to the same stream.

Well, mostly to the same stream, and then lay their eggs. Now, what is this? Is this migration, or is it not migration? If you ask an ecologist, they are going to say that, look, the individuals have moved, you know, thousands of kilometers. Obviously, it is migration. But if you ask an evolutionary biologist, they are going to say that, look, If we define the population as the fish of a particular stream, the same fish are going out and then coming back. So, in terms of the population, there is no influx of new alleles and therefore, evolutionarily speaking, we cannot call it a migration.

So, this is, you know, the two definitions are absolutely consistent in their own place. The problem happens when you know one person uses migration in one sense and the other person uses it in a different sense. So, to ameliorate this problem, many textbooks use the term gene flow instead of migration. Frankly speaking, it is a better term. However, in many textbooks, particularly the slightly older ones, The term migration is used much more commonly, and that is the one that we are going to use in the context of our discussion as well.

Now, in order to figure out what exactly migration does and in what way it affects us, it is probably a good idea to take a very Create a toy model in our mind, see what it does in the toy model, and then extrapolate it into more realistic scenarios. So, the toy model that we have in our case is what is known as the continent-island model. So, in this model, you assume that there is a huge piece of land or water, whatever you call it, the continent. And there is a smaller population that you call the island. Let us assume that both are exchanging individuals with each other.

But the number of individuals, N, that is being exchanged is a constant. Therefore, although there is migration, the population size is not changing either on the continent or on the island. So, then in this context, you explicitly assume a 1 locus 2 allele scenario And let us assume that for the allele A1, the frequency on the continent is Pc and the frequency on the island is Pi. The third major thing to assume is that the size of the continent I mean the population size of the continent is much larger than that of the island. And as a result, even though migration is happening to and fro.

The allele frequency Pc on the continent remains constant, whereas it is only the allele frequency on the island. which keeps changing and all other conditions of Hardy Weinberg are valid Which means there is no selection, there is no mutation, there is no drift, etcetera, etcetera, etcetera. So, now the question is that we have to think about migration, right? Now there is a slight bit of, well, I would not call it a trick; it is just a slightly different way of thinking about migration. In this particular context and for this model, the way we think about migration is that, you know, let there be a population. Of constant size in the population, you know the island where the migrants make up a fraction m in every generation.

This is the key phrase: migrants make up a fraction, m. So, just to give you a numerical example. Suppose there are 100 individuals in our island focal population, and let us say 25 individuals have come from outside. So the total number of individuals is $100 + 25 = 125$, and out of this 125, 25 are migrants, which means that $m = 25 / 125 = 1 / 5 = 0.2$. So this is the only thing that you have to keep in mind: the migration rate. It refers to what fraction of the total native plus migrant population the migrants represent. So m is equal to, you know, the number of migrants divided by the total size of the population. The common mistake that students make when given a situation like this is That they assume there are 100 individuals in the population; 25 individuals came in. Therefore, the fraction of migrants is 25/100, which is 1/4, which obviously is wrong.

So, if you do not make that mistake, everything else in this context is actually rather simple. So, with this background, we are now going to ask the question: how will migration affect the allele frequency on the island? And in order to do that, we will go to our simulator. Here it is, and we will start the tool. We will start the simulations; we will do simple individual simulations. So, remember that all other conditions of Hardy-Weinberg are operative; therefore, the population size is infinite.

Number of generations, let us make it 1000, and the starting frequency of allele A1; this is in the context of the island, so this is Pi. So let us keep it at 0.5, does not matter and additional settings click on migration, migration rate, So let us say we will keep the migration rate at 0.25, and the migrant allele frequency. So this is what we call Pc. So we will keep that. Let us say the other one we have kept at 0.5; let us say we will keep it at 0.9. So our allele is more frequent on the continent than on the island.

Excellent. So now let us run the simulation. We run the simulation, and what do we see? So we see that the frequency of the A1 allele starts from 0.5, and, Then, it rapidly climbs and basically stabilizes at about 0.9, which is the larger of the two frequencies. So this is also the frequency that was present in the continent.

So what is happening? Is it settling on the larger of the two frequencies, or is it settling on the frequency of the continent? How will you figure that out? So the only way to or the simplest way to do it is that now we invert the situation. So we now set the starting frequency of allele A1 at 0.9 and the migraine allele frequency at 0.5. If it is settling on the larger one, then it will still go to 0.9. Whereas if it is settling on Pc, the migraine allele frequency should start from 0.9 but go to 0.5. So, let us see what happens. We run the simulation, and there you go. It starts at 0.9 and very quickly settles down to 0.5. So, it looks like it does not matter whether the continent's allele frequency is higher or lower. Whatever the continent's allele frequency is, it will settle at that value. You can play with this, you know, with a few more values, and you can actually satisfy yourself that this is what happens.

So, until this point, we kept the migration rate at 0.25. Will things change if we change the migration rate? So, let us set the migration rate; let us say we make it 0.35. We are just playing here; any value is fine.

So, we make it 0.35 and let us see what happens. Here, it takes about 50 to 60 generations to hit 0.5. So, now we do this again, and it again settles at 0.5, except that now it is settling slightly quicker; it is taking about 30 generations or so.

So, now let us make it 0.5; let us make it even higher. What happens? It again settles at 0.5, except that now it seems to be doing so even faster; it is doing it in about 20 to 21 generations or so. So, basically, what happens is that if you increase the migration rate, Then the behavior does not really change; it still settles on whatever the value is at the continent. But what happens is that as the migration rate goes up, it approaches that equilibrium value, the continental value, faster and faster.

So, of course, you can play with this on your own and satisfy yourself as to whether this happens every time or not. So, in the meantime, we are going to go back to our slideshow, and here are the two observations that we had. Irrespective of the starting frequency on the island, the continent, or the migration rate, The equilibrium frequency on the island is whatever the frequency is on the continent. The rate at which the continental frequency is reached, the equilibrium frequency, depends on the migration rate. The higher the migration rate, the faster it reaches there.

So, now, of course, one has to ask the question: why exactly is this happening? And in order to understand why this is happening, We will do some very quick, simple algebra to derive the relationship between allele frequency and migration. And as I said, the continent-island model, just to refresh your memory. So, we have the continent where the frequency of allele A1 is Pc; we have the island where the frequency of allele A1 is Pi. In a one locus two allele scenario where Pc is a constant, the number of migrants exchanged in both directions is N. Therefore, you know the population size does not change and all other Hardy-Weinberg conditions are applicable.

So, there is no selection, there is no drift, there is no mutation—nothing; the only thing that is happening is migration, right? So, if this is the case, what exactly happens to the frequency of allele A1 on the island after one generation of migration? So, let us assume that we call the frequency of allele A1 on the island Pi after migration; that is what that prime tells you. As now that migration has happened on the island, you have the migrants and the natives, right? So, the migrants have a certain frequency of allele A1, and the natives have a certain frequency of allele A1, right? So, the migrants we have already mentioned will have a frequency of Pc, and the frequency among the natives will be Pi. Now, we know that the resultant frequency is simply going to be these two frequencies. The sum of these two is weighted by their respective fraction in the population. So, what is the fraction of the migrants in the population? By our very definition, that fraction is m.

So, this is $m.Pc + Pi(1 - m)$. What is the fraction of the natives? This is m. So, this is $(1 - m)$. So, now you realize why it was defined in that particular way. The math becomes extremely easy in this step.

So, $Pi' = m.Pc + Pi(1 - m) = m.Pc + Pi - m.Pi = Pi + m(Pc - Pi)$; we will take the m outside. So, this is the amount by which the allele frequency on the island is going to change. Now, ΔP which is sorry ΔP is the amount by which the change is happening. So, $\Delta P=Pi'-Pi = Pi+m(Pc-Pi)-Pi$. So, of course, this and this gets cancelled, $\Delta P=m(Pc-Pi)$.

So, this straight away tells us that the magnitude by which ΔPi is going to change on the island is dependent. On the multiplication of the migration fraction and the difference between the continental frequency and the island frequency. So, obviously, the larger m is, the faster you know the greater ΔPi is going to be. And the greater ΔPi is, the faster the island population will change. So, that is why if you remember, as we increased the value of m, the whole population's allele frequency change faster and faster and go faster to the equilibrium. Now, what exactly happens at equilibrium? At equilibrium, of course, $\Delta Pi =$

0, which means that this entire thing implies that m(Pc - Pi) = 0. Now, this is going to happen in two ways: one way is if your migration rate itself becomes equal to 0. But if your migration rate becomes equal to 0, then essentially you have two populations which Each one of them is in Hardy-Weinberg equilibrium because all other conditions of Hardy-Weinberg are applicable. So, obviously, at that point, there will be no further change, which is a very trivial scenario.

But if the migration rate is not 0, and if migration is continuously happening, then the only way in which your ΔPi can be equal to 0 is When Pc becomes equal to Pi, right, (Pc-Pi)=0, which means Pc=Pi. In other words, at equilibrium, this is the only condition possible. Now, of course, Pc can become equal to Pi in two ways: either Pi goes towards Pc, or Pc goes towards Pi, right? But we have started by assuming that Pc is a constant. Therefore, the only thing that can happen in this context is Pi goes toward Pc, which is exactly what we were seeing irrespective of what the island's frequency was. And what the continent's frequency was after some time, the island's frequency settled on the continent's frequency.

In other words, if you are having continuous migration like this, then After some time, both populations are going to have the same allelic frequency. This is a way in which you know it is a homogenizing factor, but the homogenization is not in terms of a loss of an allele. Or something, what is happening is that the two frequencies of the two populations are becoming equal. Now, as we saw, this is obviously a toy case, right? But it is very easy to go. From this toy case to a slightly less toy case, and that is the migration model to patch model, right? So instead of a continent and an island, now we have patch 1 and patch 2, or you can even call them Population 1 and Population 2. So the only thing we assume here is that the number of individuals being exchanged. That number of individuals remains the same in both directions, which essentially means that The m is also the same for both populations, so both populations start with the same population size. They are exchanging the same number of individuals, and therefore, the population size in the whole remains constant. So if you have a situation like this, which, as you can see, is slightly better than slightly less restrictive than the previous one, Then what is going to happen? Using the same logic that we used for the continent-island model, we can say that P1', so this is now for patch 1, $P1' = (1 - m) * P1 + m * P2$, right, because fraction m is made up of the migrants. which are individuals from the second population in which the frequency of the allele is P2, so $m = P2$ The fraction $(1-m)$ is made up of the natives, among which the frequency is P1, so $(1-m).P1$. By the same token, P2', which is the frequency after migration, is $P2' = (1 - m)P2 + mP1$. Now, we need to figure out what the overall frequency of the allele A1 is. Taking two populations together, in some sense, forms the metapopulation. So, that thing, the overall frequency, I am just going to call that P; this total, in some sense, is P' after this. $P' = (P1' + P2') / 2$; right, why the population size is exactly the same is that the same number of individuals are being exchanged.

So, it is simply the average of the two: $(P1' + P2')/2 = [(1 - m)P1 + mP2 + (1 - m)P2 + mP1]/2 = [P1 - mP1 + mP2 + P2 - mP2 + mP1]/2$. Now, you can already see I have a $-mP1$ here and a $+mP1$ here; they cancel each other out. I have a $+mP2$ here and a $-mP2$ here; they cancel each other out. Which is simply $P' = (P1 + P2) / 2$; this is, you know, the previous generation's thing, right? So, this is essentially the average allele frequency in the two patches before migration. So, the overall allele frequency after migration is equal to the overall allele frequency before migration.

Now, this is not at all surprising; why that is, because just think about it intuitively: forget about all the math; what is happening? we are simply taking N individuals from one population and pushing it here and another N individuals from this population and Pushing it here, right? Nobody is getting killed; there is no sampling, there is no drift, no mutation is happening, no selection, nothing. So, all we are doing is redistributing the individuals from this population to that population, and therefore, very obviously, The overall frequency in the population is going to remain constant. Which is precisely what you are going to see over here: $(P1+P2)/2$. Now, what exactly is happening in either of the two populations? So, as we said, P1' dash is this thing.

So $\Delta P1 = P1' - P1 = (1 - m)P1 + mP2 - P1 = P1 - mP1 + mP2 - P1$. So I am just taking it from here, right? So this dash is not there. So this and this get canceled; I will take the m outside; $\Delta P1 = m(P2 - P1)$, right? Now, similarly, this is delta $\Delta P1$; similarly, $\Delta P2 = m(P1 - P2)$. Now, again, take any one of them; let us take the first one, alright. So, at equilibrium, what is going to happen is that $\Delta P1=0$, and for that to happen, Either we will have to face the situation when there is no migration, which, as I said, is the trivial case. The other situation in which it can happen is if $P1 = P2$; right, you get the same condition from here: $P1 = P2$.

In other words, if you have two populations that are exchanging migrants, then the entire thing is going to stabilize. When the two allele frequencies on the two patches become equal, you know the populations. Which, if you remember, is precisely what we saw in the continent-island model. So, this is the main thing here, the main insight that when you have migration, Then, the allele frequencies tend to become homogenized over space and This is a very, very powerful way in which, you know, it ends up affecting evolution. In fact, this migration thing ends up interacting with all the other evolutionary forces that we have seen till now.

But those interactions and real-life examples for those are what we are going to see in our next discussion. See you then. Bye.