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 26

Genetic drift: examples and features

Hi, so in our last discussion, we did a bunch of numerical simulations to understand a few features of genetic drift. And I promised you that in the next one, which is this one, we are going to look at some examples and a few features. So, before we get into that, if you remember, I made a statement that was In the context of drift, the probability of fixation of an allele due to drift alone is equal to its starting frequency. And we had done some simulations, and we had seen that this generally looked to be a correct statement. So, today before we get into the examples, I will give you the demonstration of why this property is true. It is a very simple thing; it is just a, you know, bunch of logical statements connected to each other.

So, assume that we have a diploid population that has n individuals, and in this particular population, Let us assume that drift is the only evolutionary mechanism operating. So, there are no mutations, there is no migration, there is no selection. Now, since the diploid population is of size N, the total number of alleles in the population has to be equal to 2N. Now, let us make the assumption that each one of these 2N alleles is different from the others.

So, each one is a unique allele. So, if that is the case, what will be the frequency of each allele? Obviously, the frequency of any given allele is going to be 1/(2N). Now, when we say that drift is operating, by definition, drift is unbiased, right? So, since drift is

unbiased, all these 2n alleles, each one of which has an equal probability of getting fixed. If that is the case, what is that probability going to be equal to? That probability is going to be simply 1/2N. Now, let us assume that all these alleles are the same as each other.

Let us assume that there are two kinds of alleles, A1 and A2, such that there are x copies of A1 and y copies of A2. So, x + y = 2N. Now, if there are x copies of A1, then what is the overall probability that a copy of A1 will get fixed? That is simply going to be x*(1/2N). Now, if that is x*(1/2N), then also note that x*(1/2N) is the starting frequency of A1. Therefore, we have proven that the probability of fixation of an allele due to drift alone, and nothing else, is going to be equal to It's a simple logical thing, nothing more than that.

So, now we get on with our examples. Now, when it comes to drift, there was a very important, very famous experiment done in 1956 by a person named Buri. And this is an absolute textbook example; every single textbook of evolutionary biology has A very high probability of having this experiment; so, we will spend a little bit of time understanding this one. So, this is an experiment using Drosophila melanogaster, and what Buri did was found. 107 Drosophila populations, each with 8 pairs, which means 8 males and 8 females.

Then these populations were not founded at random; they had a property in that every single founder, So, 107 flies into 16, all of them were heterozygous at a locus called the brown locus or the bw locus. Now, at this locus, the bw gene codes for an eye color gene. And therefore, the beauty of the experiment is that all three genotypes are discernible only by looking at the phenotype. So, you look at the eye color and you know that if it is a white eye color, then it is bw/bw homozygous; if it is a Bright red-orange eye color indicates that it is a bw75/bw75 homozygote, and if it is a light orange color, then it is a bw/bw75 heterozygote. So, what he does is start the entire experiment with the heterozygote.

So, everybody has light orange eyes. Now, the very important reason for choosing this

particular gene and nothing else is that there is no known fitness effect of these two alleles. Why is this important? Because remember, we are studying drift alone, and therefore, we do not want selection to mess things up. And all he did in the experiment was count the number of genotypes in the population for every single generation. And every single generation, he again collected 8 pairs, 8 males, and 8 females at random and used them to found the next generation.

So, basically, the eggs laid by these pairs—those are the ones that created the next generation—did this for 19 generations. So, here is what the experimental data looks like. Now, I will tell you exactly how this graph is drawn. So, each panel that you see is for a generation. You can see it says initial population, generation 1, generation 5, etc all the way up to generation 19. Now, what you have on the x-axis is the number of BW75 alleles, and what you have On the y-axis is how many populations had that particular number of BW75 alleles, so it is a kind of histogram. Now, remember that in every generation, the population size is kept constant at 16: 8 males and 8 females. So, since there are 16 individuals, you are going to get 32 alleles; that is all you can have. Therefore, the number of bw75 alleles can vary from 0 to 32; that is the range.

And every generation, he is simply counting how many populations have zero bw75. How many have 1, how many have 2, how many have 3, all the way up to 32? So, let me walk you through this panel slowly. So, let us look at the first one, the first panel over here. Now, if you look at this panel, you see that in generation 0, all the populations have 16 bw75 alleles. Why? Because all of them are heterozygotes, all 107 have 16 bw75.

What happens in generation 1? Drift has started operating because the population size is very low, and now the populations have different numbers. bw75 alleles, and you can straight away see in the second panel that the numbers have deviated from the value of 16. There are some that have about 14, some that have about 22 or 24, and so on and so forth. You go to generation 5, you go to generation 10, and now you can see that the number of, you know, bw75 alleles. The population is varying a lot, which is why you know you are seeing the histograms getting spread out.

And generation 10 is when the populations start getting fixed for either the bw allele or the bw75 allele. How do I know that? Because generation 10 is when, for the first time, we are getting. You know some populations that have 0 or some populations that have 32. And then, as time progresses, generation 15 and generation 19, you can see that the bars on the two sides are becoming. Taller and taller, and the ones in the, you know, middle between these two extremes are becoming smaller and smaller, right? Why is that happening? Because as time progresses, more and more populations are losing the bw75 allele or getting fixed for it.

So, by generation 19, as you can see, 30 populations completely lost the bw75 allele, whereas 28 populations. were completely fixed, which means that every individual had only the bw75 allele and nothing else; 28 were of that type. So, the reason that this particular experiment is so famous is that it is very nicely conducted. It shows you the two major properties of drift that we observed from numerical simulations. One is that, due to drift alone, if you start with replicate populations, then All of which have the same genetic makeup due to drift alone; over time, these populations can diverge from each other.

So, that is exactly what you are seeing: you are starting with, you know, just one tall building and how. That thing is now being distributed over time across a various number of BW75 values. And the second major thing it shows is that the ultimate fate of drift is either fixation or loss, which is what is happening here. By generation 19, you can see that 58 out of the initial 107 populations, which is about 54.2 percent, Slightly more than 50 percent has become fixed in one direction or the other.

So, we will get back to this particular experiment in our next discussion where we are going to, you know. When we look at how to measure drift, we will see what the measurement of drift is in this particular case. However, before we get there regarding the measurement of drift, we need to talk about two special cases of genetic drift. In these two special cases, the first one is called the founder effect. What is the founder effect? So

suppose you have a population that has some genetic composition, some allele frequencies.

And then from this population, a small subset comes out and goes to settle somewhere else. And because it is a small subset, there is always a probability that the genetic and allelic frequencies This subset is very different from the allelic frequency in the main population. And now, when this small subset goes and settles in a new place, and then the population size expands over there, then, After some time, that population is going to have a very different genetic composition compared to the ancestral population. But that need not necessarily have anything to do with selection or any matters that might simply be because of the founders. Who began that new population did so due to sampling error and had a very, very different allelic composition to begin with.

So, just to give you an example, you know, a pictorial example, suppose you have a situation like this where you have, let us say, Blue circles and red squares at some frequency, and let us say we had three different founding events from this. In one case, just due to sampling alone, you only have red squares founding that population. In another case, again, due to Sampling alone, you have blue circles founding the population, and in the third case, you have again some red and some blue. Now, after some time, if you are not aware of the fact that the founding happened this way, you are going to Look at these populations and say, "Aha, look at this top one; it is fixed for red.

That means that..." The red allele is providing a major selective advantage here, which is why this population is fixed for red. Or similarly, you know it is fixed for blue, and you know the same logic. However, the real reason this has happened is something much simpler: it is because when they started off. They started with a different frequency, you know, all reds or all blues, and that is why they are all red and all blue today. The second, very interesting case is actually just a variant of this, which is known as the bottleneck effect.

So, in the case of the founder effect, you have one population, you know. Being the

ancestral population, and from that, a small subset going somewhere else. In the bottleneck effect, nobody goes anywhere; you have the large ancestral population, and due to some reason, Some environmental catastrophe or some political stuff, whatever it may be, the ancestral population's size becomes very small. And when the size becomes very small, it is entirely possible that the Allelic frequency changes again due to sampling error and nothing more. So, just to give you an example, suppose you think about the ancestral population as a bottle full of Three different colors of marbles: green ones, orange ones, deep orange ones, and light orange ones.

And let us assume that the population size has shrunk, which is represented by the narrow neck of the bottle. And now you know that if you shake the bottle a bit, some marbles come out, some bits come out, and that represents the fact. That a small subset is what is left due to your environmental catastrophe or whatever. And the surviving population now has a very different composition, which is what you see over here. The bright orange marbles or the beads have been completely lost.

Now the population only consists of green and light orange ones. So, again, if you look at it and you are not aware of the fact that this kind of constriction has happened. A bottleneck has occurred; you are going to think, "Aha, the bright orange allele had some major fitness-related issues." It reduced the survivorship of fecundity significantly; therefore, it has been thrown out of the population. But in reality, that is not what has happened; all that has happened is, again, a kind of sampling error.

Now, there are a few points that I need to make in the context of the founder effect and the bottleneck effect. I have been sitting in interviews at IISER Pune for the last 17 or 18 years, and one of the common questions What we ask people is, "What is genetic drift?" Many students end up saying, Oh, genetic drift means the founder effect and the bottleneck effect. That is not really true. As I am showing you, the founder effect and bottleneck effect are two special cases of genetic drift. Sampling errors occur in two particular ways, but these are not the only two ways in which genetic drift can occur.

As we saw in our simulation, genetic drift can happen even with constant population sizes, as long as that population size is finite. And therefore, genetic drift needs to be defined as those random changes in allele frequency and not as the Founder or Bottleneck effect. This is a very important message, you know, that I need to give to all students who appreciate the correct definition of drift. The second point is that, as I said, under the hood, all that is happening is a sampling error, and the sampling error is happening in these two cases because you have a large population; from that large population, you suddenly have a A very small subset, and that subset may not be representative in terms of its allele frequencies of the main population. Therefore, although people tend to differentiate between them in terms of genetic signatures.

They are often the very same, and until you have special information, it becomes very difficult to tease apart. Whether a particular signature that you are seeing is due to founder effect or a bottleneck effect. And finally, the main thing that happens over here is that the population genetic composition The offspring generation is very different from the parental generation. So, because of the sudden reduction in population size, the main thing that ends up happening is Typically, there is a reduction in the genetic variation, which is the main thing. The founder effect and bottleneck effect typically lead to a reduction in genetic variation.

I mean, so does genetic drift, but the founder effect and bottleneck effect are very, you know, stark. And every one-step big ticket effect that you can see. So, let us look at a few examples of both the founder and bottleneck effects. So, this is one of the very famous, big examples of founder effects. So, we know that in humans, Homo sapiens, we originated somewhere in East Africa.

And then, from East Africa, we actually migrated to different places. And these migrations did not happen in one, you know, event; they happened subsequently. So, we went from point A to point B, and then after some time, one group went from point B to point C. Another group went to point D, and so on and so forth. So, as far as scientists have been able to trace, human beings originated.

You can see the yellow star here that is in Addis Ababa, somewhere in East Africa. And then from there, there were lots of migration events that happened in other parts of Africa. These are what is represented as these red inverted triangles. Then, though there was one migration event that happened towards You know the Middle East is going towards the north of Europe and Russia and all that. And then, from, you know, the Middle Eastern population, one set went towards Western Europe.

Another set went towards East Asia. And then from there, from China and other places, there were other migrations. That happened in the northeast of Russia, eastern China, Japan, etcetera. And, very importantly, this Indian, sorry, the Middle East group sent out one branch toward far East Asia. And from there, they went towards, you know, Indonesia, the Philippines, all those places. And from there, they went into North America, and from North America, they went into South America.

So, this is how we have been able to piece together the migration routes of ancient humans over time. Now, appreciate one thing: each time a group goes from one place to another, you are going to encounter a founder event. So, you know, for example, when they go from Africa to the Middle East, that is one founder event. When they go from the Middle East to, let's say, Western Europe, that is another founder event, and so on and so forth. So, if that is the case, what we have is a series of founder events, and what I told you is that.

Each time a founder event or a bottleneck event happens, typically the genetic variation goes down, right? So, what people did was that these people, you know, Liu et al., ended up looking at the genetic variation. In people from these various parts of the world, they ended up showing that as you go further away. From East Africa, in terms of migration, the genetic variation present in humans continues to decrease. So, the maximum genetic variation is actually found among human beings in Africa.

Then the second largest you find among people who are in the Middle East from there.

You know the amount for people who came to India and other parts goes down. And then when you look at people who went to Far East Asia, these are the red squares on this; it goes down further. And then when you look at those individuals who went into, you know, Indonesia, the Philippines, etcetera; their genetic diversity goes down further. And then if you look at the individuals, you know what is shown on this map in white circles; these are the guys.

Those who are in North America and South America are the so-called last in terms of the founder events. If you look at their mean heterozygosity, I will tell you what mean heterozygosity is. It is a measure of how much variation there is in the population that you know. If you look at that, you see that it decreases further. So, in other words, you see a very nice negative correlation between how far you are.

From the center of origin of humanity, what is the level of genetic diversity that you have? So, this is a very nice example of what you know continuous founder events can do. Now, one can, of course, ask that you know this might have been due to selection. How do we know that this is not due to selection; this is due to drift alone? That is because of the particular genetic sequences they are checking. The variation, as far as we know, consists of neutral sequences. In other words, sequences that do not have any effect on fitness are therefore not seen by selection.

Hence, we know that this is due to drift alone. This is a big ticket for all of humanity taken together. Let us look at a more specific example of the founder effect in humans. Now, to people in India, this gentleman, I am sure, needs no introduction: Bollywood actor Hrithik Roshan. Now, many of you might be aware that Hrithik Roshan has a condition known as polydactyly.

He has an extra finger here. Now, it turns out that polydactyly is actually pretty rare among humans, and its normal frequency is about 4 to 12 per 10,000 births. However, there exists a population in the US where the frequency of this particular condition, Polydactyly is much, much greater, and how much greater? This is about 7 per 100 births.

So, this population is the so-called Amish population. So, the Amish are a known migrant community. So, they migrated from Switzerland, in particular, to the US somewhere in the early 1700s.

So, these are fascinating people because they believe in, you know, staying away from modernity. So, in many ways, these people are retaining many of their cultural practices from, you know, 300 years ago. They typically do not use any machines, you know; most of them believe in very hard work. Extremely God-fearing people, and most importantly, they do not marry outside. So, the Amish marry other Amish, and because of all this, the genetics of the Amish population are absolutely fantastic.

And there are lots and lots of studies on the genetic properties of these people. But as I said, the one thing that is very interesting among the Amish is that they have a very high frequency of polydactyly. This particular thing actually comes from a single mutation in a gene known as the EVC gene. Now, why exactly is it that the Amish have such a high frequency of polydactyly? So, one of the things that the Amish do as part of their culture is maintain very detailed genealogies. Therefore, we essentially know that you know who the father of whom is.

And who is the grandfather, and who is the great-grandfather, etcetera, etcetera? So, scientists have actually traced the entire genealogy of these Amish migrants, and they have figured out that in 1744, There was one couple, Mr. and Mrs. Samuel King, who came from Germany to Pennsylvania.

And it is not clear whether Mr. King had it or Mrs. King had it. But it is pretty clear that one of them had this particular mutation; it is a resistive mutation. And after that, this one mutation essentially ended up spreading among the Pennsylvanian Amish. And the entire problem that we have today, 7 per 100 births, is because of that one couple. In other words, if those two people, the husband and wife, had missed their boats or if the boat had sunk, you know. God forbid during the, you know, transatlantic cruise, then the Amish people would not have had this particular condition.

Now, I am talking in terms of polydactyly, but actually, what they have is a bigger condition than that; it is a condition called EVC. And because of that, there are many other physiological problems that the Emish are saddled with, and as I said, all because of. Mr. and Mrs. King came in 1744, a massive and very well-known example of founder effect in humans.

So, now we go to bottleneck effects, and for bottleneck effects, we are going to Micronesia. Now, I am sure many of you have not heard about Micronesia. So, Micronesia is, you know, a bunch of islands located somewhere north of the Philippines. and very small tiny islands, and amongst these islands, you know, that constitute this nation of Micronesia. There is one known as Pingelap, and I am showing you a satellite picture of Pingelap just to tell you you know how small it is; you are not getting an idea of how small it is, but you can see that somewhere here. You know, somewhere close to the center, you can see a strip—a white strip—over here. This is a runway just for scale; this is a runway. So, this tells you how small this entire place is. Now, why is this place interesting? Why is this place so well known in the whole of biology? That is because there is a condition known as achromatopsia. What is it? It is an eye disease that leads to complete color blindness. Now, what do I mean by that? See, normally when you have color blindness, it is about the inability to distinguish between certain pairs of colors. So, let us say that in red-green color blindness or blue-yellow color blindness, people cannot distinguish between those pairs. In the context of achromatopsia, they have complete color blindness, which means They cannot distinguish any color; they see the world in black and white, and that is bad. And more importantly, if it were just seeing in black and white, it would be bad, but in reality, it is way worse than that. These people, their visual acuity, which means that how sharply they see something is really, really bad. And, more importantly, they cannot go out in sunlight. The moment they go out in bright sunlight, which is common in tropical places, the sun is obviously really bright there. The moment they go out in the sunlight, their eyes start watering and hurting like anything. So obviously this is something, and if you are on an island, this is not like a highly developed kind of society. Where you have lots of ways to, you know, not go into the

sun, etcetera. These are, you know, economies that depend on nature and therefore, Not going out into nature during the entire daytime is a big problem for these people.

So, now obviously the question is, how did it come to be? So, it turns out that achromatopsia is known outside of Pingelap as well. However, if you look outside, the prevalence is very small; 1 in 20,000 or 1 in 50,000 is the worldwide rate. However, if you look at Pingelap Island, for those individuals, 6 to 10 percent of the population has this condition. And about 30 percent are carriers. Now, how come something that is so problematic for the species is so widespread? And it turns out the story over here is fantastic, or not for them, but from a scientific point of view.

So, it turns out that in 1775 there was a very, very massive typhoon called Liengkieki. I hope I am pronouncing it right, Liengkieki. Now, due to this typhoon, about 90 percent of the population of that entire island was wiped out; they just died. And after the typhoon was over, all the vegetation and everything were affected. So, there was basically no food for a very long time; there was almost a famine-like condition.

A huge number of people died due to starvation after that. So, finally, after this entire disaster you know was over, there were only about 20 people left on that island. And the current population of Pingelap, which is about 250, is made up of all descendants of these survivors. 250 is the population now because many people have migrated outside Pingelap. But the population had actually gone up much higher; I think it reached 750 to 800 at some point.

And even every one of those was actually a descendant of these survivors. Turns out that one of the people who survived was the local ruler, and it also turns out that This person happened to be a carrier of a recessive mutation for a gene known as the CNGB3 gene. And this particular gene encodes one component of a protein that is crucial to the functioning of the cone cells in our eyes. So, you know that we have rod cells and cone cells, and you know that. Cone cells in our eyes are the ones that are important for color vision.

So, the reason they are important for color vision is because of proteins. One of the components of that protein is the gene coding for it that gets a mutation. That is it. Because this person survived and because he was the ruler, he probably ended up having a larger number of babies. Relatively speaking, the condition has spread, and today the Pingelapese population is suffering from this kind of situation.

So, just for fun, I want you to watch this particular short video. It is just a 2-3 minute video that takes you to the Pingelap Atoll. It is absolutely beautiful; I mean, natural beauty is superb, but it focuses on the condition. And how painful it is for the people of that particular place. Note that I did spot a few errors, you know, in the facts presented in the video. So, if you compare the facts stated in the video with what I am telling you over here, you will see some discrepancies.

I think mine are correct because I ended up cross-checking my facts with multiple papers. I do not know whether they have done any of that. The second thing that I want you to look at is this very famous book. In fact, it is this book that made the Pingelap you know condition famous all over the world. This is a book called "The Island of the Color Blind" by Oliver Sacks.

A brilliant author has written on multiple topics related to evolution and many other subjects. This particular book I have read, you know, excerpts of it are absolutely brilliantly written; highly recommended. And finally, if you want to learn about genetic drift, then the Amoeba Sisters, You know, the version of genetic drift in this particular video is really, really good. The information content is roughly similar to what we have discussed so far, but obviously the way they present it is very, very fun. So, until this point, we were just talking about drift, you know, changing the composition, but at some point, We have to be able to measure the effects of drift, because without measuring, We cannot really figure out what exactly it is going to do to evolution. And how to measure drift and what those measures will lead to, those are the things that we are going to do in the next discussion. See you, bye.