Selection

Evolutionary ecology

Up until this point, we have not consider the role of evolution. That is, we have optimized reproductive output given fixed reproductive value of different classes and fixed parameters of birth rate, fecundity, and survival. In this chapter, we'll allow demographic rates to evolve in accordance to selective pressures. By incorporating evolution, we will introduce different terminology and, in some sense, different goals. The goal of population ecology is largely to model changes in population size over time. The goal of evolutionary ecology does not necessarily care about population size, but about allele frequencies. These two goals can be merged by considering selective forces on population growth rates. Where population ecology deals with

$$N_{t+1} = f(p_t, N_t)$$

where p_t is gene frequency at time t, evolutionary ecology deals more with how changes in p_t are a function of population size

[p_{t+1} = g(p_t, N_t)]

If we were able to model $f()$ and $g()$, we would be able not only to predict population dynamics, but also changes in allele frequencies over time, and how both vary as a function of population density.

Darwinian fitness is the underlying 'currency' of natural selection, in that natural selection will favor certain genotypes (not individuals) and this selection will have an underlying basis in the contribution of those genotypes to future generations (linking population dynamics to genotypes and selective pressures). We might even not include the role of genotypic variation directly into our population dynamic models as we have covered so far. That is, genotypes may vary in birth rate or survival (something that we would track in our population models), but also could vary in something like access to mates, which may tied to birth rate or reproduction probability (i.e., we haven't modeled this explicitly, but some of this variation may be subsumed in other demographic model paramters).

What is perhaps interesting (or frustrating) about fitness is that we do not care about absolute fitness, but only *relative* fitness. That is, a certain genotype may be favorable in one population, but that same genotype may no be favored in another population. To put this into tangible terms, let us consider a simple case of two alleles (a and A) in a diploid outbreeding sexual species. This creates the possibilities of offspring to be aa, Aa, or AA. Each one of these potential genotypes has a fitness value λ_g, where g can take forms of aa, Aa, or AA. This absolute fitness λ_g can take a variety of different forms, dependent on the allele frequency g and population size N.

((go over Figure 9.1 in terms of how absolute fitness of a genotype may vary as a function of population density and gene frequency))

The case of no frequency or density dependence

This corresponds to the situation in Figure 9.1a. We will start by setting up this simple model, building on it to develop more complex scenarios. That is, we haven't tracked allele frequencies to this point, but now we have to. Let's consider a semelparous life hisotry with discrete nonoverlapping generations. The species has a sexual diploid life cycle, and adults die after they reproduce. Selection can broken into many different components: mating success, birth rates, gamete survival, and so on. The point of the underlying model is to reflect reality without being overly paramterized or failing to capture the system. One thing that we *need* to incoporate that should be immediately clear is that we need to consider variation in survival or birth rate

as a function of genotype. Doing this creates the variation in genotype growth rate λ_g that is the basis for selection. For instance, we can consider the growth rate of the AA genotype as

$$\lambda_{AA} = s_{AA} b_{AA}$$

where genotype AA has some genotype-specific survival rate (s_{AA}) and birth rate (b_{AA}). What we care about is the frequency of genotypes in the population, so we will track these in our simple example in the following way, where we consider the frequency of gene A to be p and the frequency of a to be 1-p, which we will call q.

Genotype	Frequency
AA	p^2
Aa	pq
aA	qp
aa	q^2

It is important to note that the above assumes completely random mating (i.e., the genotype of an individual does not lead to assortative mating, so individuals can't pick mates based on genotypes). This also requires the population to be infinitely large. The above table corresponds to what is normally referred to as a the *Hardy-Weinberg equilibrium*. It makes some assumptions, some of which we'll address below and some of which we'll cover before we leave this section.

We can downscale the number of individuals that we might expect to be of differnt genotypes by simply multiplying the allele frequencies by N_t, so at time t, the frequency of genotypes would be AA ($p_t^2 N_t$), Aa ($2p_t q_t N_t$), and aa ($q_t^2 N_t$). How do we know the frequency at time $t+1$? We might be tempted to simply multiply the population of each genotype by the genotype-specific growth rate (e.g., λ_{AA}). Why is this wrong?

We need to consider the individual alleles separately, as A can form Aa or AA. So let us consider the situation where we simply want to model the presence of allele A. Then the frequency of A alleles is the number of A alleles after selection divided by the number of all alleles after selection. In mathematical terms, this would correspond to

$$p_{t,s} = \frac{2p_t^2 N_t \lambda_{AA} + 2p_t q_t N_t \lambda_{Aa}}{2p_t^2 N_t \lambda_{AA} + 4p_t q_t N_t \lambda_{Aa} + 2q_t^2 N_t \lambda_{aa}}$$

or the simplied form

$$p_{t,s} = \frac{p_t^2 \lambda_{AA} + p_t q_t \lambda_{Aa}}{p_t^2 \lambda_{AA} + 2p_t q_t \lambda_{Aa} + q_t^2 \lambda_{aa}}$$

This simplification has removed N_t from the equation. Also note that the denominator of this equation is the average absolute fitness across all the genotypes, which we can refer to simply as $\bar{\lambda}_t$, where the overline indicates the mean value and the subscript t makes it clear that this value will change as allele frequencies change. We can use the above equations to estimate allele frequency of a particular allele. For instance, if we wanted to know the frequency of A at time $t+1$, it would be

$$p_{t+1} = p_{t,s} = \frac{p_t^2 \lambda_{AA} + p_t q_t \lambda_{Aa}}{\bar{\lambda}_t} \tag{1}$$

under the assumption of random mating, making the absolute number of AA individuals in the population at time t,

[N_{AA, t+1} = p_{t,s}^{2} N_{t}]

This is of note because our earlier guess made without considering the absolute abundance of alleles was $p_t^2 N_t \lambda_{AA}$ (pay attention to how we defined $p_{t,s}$ relative to p_t). To revisit the Hardy-Weinberg principle, the above assumptions made above include an infinitely large population size, random mating, and no selective advantage to having a certain genotype. If we violate this, what happens?

How does natural selection affect the genetic composition of the natural population?

The above section has allowed us to track population size change and allele frequency change over time, but has assumed that there is no selective advantage of indivduals having one genotype (e.g., AA) over another (e.g., aa). Also, in the last section, we defined the overall fitness of a genotype as a constant (e.g., λ_{AA}). Most of the time though, we may want this value to change as a function of population size, or at a minimum, be compared in terms of *relative fitness*. For instance, we could use the AA genotype as the basis for comparison.

$$w_{AA} = \frac{\lambda_{AA}}{\lambda_{AA}} = 1$$

$$w_{Aa} = \frac{\lambda_{Aa}}{\lambda_{AA}}$$

$$w_{aa} = \frac{\lambda_{aa}}{\lambda_{AA}}$$

These relative fitness terms must be greater than 0, and can be substituted into the equation for $p_{t,s}$ (equation 1), replacing λ_g terms with relative fitness terms (e.g., $w_{AA}\lambda_{AA}$).

$$p_{t+1} = \frac{p_t(p_t + q_t w_{Aa})}{p_t^2 + 2p_t q_t w_{Aa} + q_t^2 w_{aa}} \tag{2}$$

Just as in the population models introduced earlier on, the frequency of allele p has equilibrium values. These equilibrium values would occur at points where $p_t = p_{t+1}$. There are at least two trivial instances of equilibria given the above equation. That is, when $p_t = 0$, and when $p_t = 1$, corresponding to instances where the dominant allele is lost or becomes fixed, respectively. But there also may be an *interior equilibrium* where p_t is still bound between 0 and 1, but is some value other than 0 or 1. We can solve for this value by dividing both sides of equation 2, assuming that $\frac{p_{t+1}}{p_t} = 1$.

$$1 = \frac{\hat{p} + \hat{q} w_{Aa}}{\hat{p^2} + 2\hat{p}\hat{q} w_{Aa} + \hat{q^2} w_{aa}}$$

where $\hat{p}$ corresponds to the equilibrial value of p. This expression can be re-arranged to

$$\hat{p} = \frac{w_{Aa} - w_{aa}}{(w_{Aa} - w_{aa}) + (w_{Aa} - 1)}$$

when $\hat{p}$ is maintained at some level not 0 or 1 (i.e., there is an internal equilibrium), we say that a *polymorphism* exists at this locus. So we have an equation for the internal equilibrium for p (i.e., $\hat{p}$). We may want to know if this equilibrium is stable or unstable, as we explored in the population models. But recall that p is simply the allele frequency of A, meaning that the actual genotype may be AA or Aa. The fitness differences between these genotypes is the driving force of the stability of $\hat{p}$, which Case explores graphically in Figure 9.3. Here, the top panel corresponds to situations where p goes to 0 or 1, respectively, while the bottom panel shows situations when the heterozygote (Aa) has the highest growth rate (stable) or when the homozygote (AA) has the highest growth rate (unstable). Explore these equilibria values with a focus on how the initial values of p affect the dynamics, and how population size is related to p (Figure 9.4).

Density-dependent natural selection

Population size N_t may complicate the dynamics of changes in allele frequencies if selection (or the relative fitness values w_{AA}) are a function of population size (density-dependent natural selection). Let's explore how this works, using the logistic model as the population dynamic framework. We previously defined a λ_g value – g corresponds to AA, Aa, and aa – where all genotypes had constant growth rates, meaning that

$$N_{t+1} = N_t \lambda_g$$

To incorporate this into the logistic model, we take the classic logistic model

$$N_{t+1} = N_t + (rN_t(1 - \frac{N_t}{K}))$$

and make it specific for each genotype (g), and make

$$\lambda_g(N_{t+1}) = 1 + R_g(1 - \frac{N_t}{K_g})$$

Given that N_t is changing at each timestep, the above equation would also suggest that λ_g changes as well. Simulating this model, we can start to explore graphically how density-dependent selection would work. For instance, consider the situation where all R_g values are equal, but the carrying capacity values K_g change (Figure 9.5). Looking at a more complicated example, where we assume that R_g and K_g are inversely related (Figure 9.6), we see how changes in population size change the landscape of p. This creates a situation where harvesting can change the allele frequencies even if harvesting is random. That is, if we reduce population size randomly, we may enter regions of that parameter space in Figure 9.5 where one genotype is favored over the other.

Harvesting in the case above is a general case of imposing some mortality cost to shift the allele frequencies. The more complicated example is if harvesting depends on the density of the species (e.g., the number removed depends on the number present). What if we were to explore a simpler example, where individuals are removed from the population as a cost of density-independent mortality (e.g., every generation 10% of the population dies).

We'll start by considering two clones (to not get confused with the genotype bits from above); clone 1 and clone 2. We will assume that both clones respond to density in the same way, so we only need to consider overall density (N_t). Solving for the logistic model equilibrium for clone 1, we have

$$R_1 \left(1 - \frac{N_t}{K_1}\right) - f_1(t) = 0$$

where f_1 is the density-independent mortality term, which only depends on time t. Solving this equation for N_t leads to

$$N_t = K_1 \left(1 - \frac{f_1(t)}{R_1}\right)$$

Since we set up clone 2 with the same population dynamics (except for clone 2 specific growth rate, carrying capcacity, and density-independent mortality terms), the above equation is the same for clone 2 with some subscripts changed from 1 to 2.

$f_x(t)$ (where x is either clone 1 or clone 2) varies over time. If we were take the average of the function, we could argue that the clone with the highest growth rate will dominate. In the case where $f(t)$ is 0, the clone with the largest carrying capacity is predicted to dominate. As $f(t)$ becomes larger, the clone-specific growth rate becomes much more important to determining the resulting competitive structure. Extending

this to a diploid sexual-reproducing species, we can simulate the model given different R_x and K_x values for each genotype (AA, Aa, and aa). This corresponds to Figure 9.9. Note from these simulations that heterozygote (Aa) is present in all situations, and is predicted to have highest fitness at intermediate levels of density-independent mortality ($f(t)$).

From these simulations, a conceptual view of the balance between R and K starts to emerge as a function of environmental-induced (density-independent) mortality. That is, in more stable environments, where $f(t)$ is low, K will determine the dominant species/clone/etc., while in more disturbed environments with high $f(t)$, a higher R is favored (corresponds to Grime's conceptualization in Figure 9.10).

Is an R-K trade-off biological reasonable?

Do R and K necessarily need to trade off, or can species be good at both? Through model simulations (Figure 9.11) and experimental evidence in protozoan communities subjected to selection and then raised at different densities (Luckinbill 1979), we see little evidence for an R-K inverse relationship (and indeed find positive covariance in simulations). Much of the current focus now has been on age-specific mortality and fecundity being the driver of life history evolution, as we touched on (without evolution) when going over age-structured population dynamics.

Density and frequency dependent competition

Above we have considered the effect of the overall N_t on the growth rate of each clone, assuming that clones have interchangable effects on one another. This is not really always the case, as competition within a species (*intraspecific* competition) tends to be stronger than competition between species (*interspecific* competition). This balance tends to be stabilizing (but Google "Chesson coexistence" if you are ready to go down a large rabbit hole on this). We can model the influence of density-dependent competition by weighting each genotype by some value ω_g, where the sum of all ω_g values is 1. This would allow us to create situations where the relative effects of one species on the dynamics of the other are dependent on the densities of competing species (or genotypes, as is this case).

$$\lambda_{AA}(N_t, p_t) = 1 + R_{AA}\left(1 - \frac{Nt(\omega_{AA}p_t^2 + \omega_{Aa}2p_tq_t + \omega_{aa}q_t^2)}{K_{AA}}\right)$$

We see how these dynamics can play out in Figure 9.12, where intraspecific limitation results in a stable polymorphic equilibrium. But this may be rare, especially if we consider the variety of shapes that density-dependent effects can take. That is, even if the ω_g values stay constant, the growth rate of the genotype λ_g is a function of population size and allele frequency. The math of this is a bit beyond the scope of what we can cover now, but a graphical representation of this fitness landscape and the role of density-dependence is in Figure 9.13. This also serves to highlight how the order of species arrival, or how the loss of a species, can result in subsequent community re-structuring.

Natural selection, adaptedness, and population density

Evolution is slippery. Just as differences in relative allele frequencies or density-dependent competition/selection can change population dynamics and evoltion, so can the *adaptedness* of species/strains/clones, where *adaptedness* is the ability of a genotype to survive and reproduce in a given environment. That is, R and K may be dependent on the environment. We will go into this in more detail in future chapters, but an example is given in the book of flour beetles (*Tribolium castaneum*). This species is cannibalistic on developing eggs, larvae, and pupae. If this cannabilistic tendency has a genetic basis, the increase in the cannibalistic genotype could lead to a decrease in overall population density, providing an instance where selection does not necessarily increase population sizes (sometimes fitness does not translate into high population density).