

## Chapter 7: Drift of a Single, Neutral Trait

© Stevan J. Arnold 2016

**Overview.-** Nonrandom sampling of parents can cause the phenotypic mean to fluctuate from generation to generation. Such fluctuations are most severe in small populations and in the absence of selection. Theory for drift in the mean is based on a model in which many replicate populations are derived from a single ancestral population and thereafter evolve independently. Such models allow us to predict the variance among replicate trait means at any generation in the future, as a function of genetic variance and effective population size. In nature, populations diverge more rapidly in the short term than predicted by drift. In the long term, populations and species diverge less than predicted by drift.

### 7.0 The concept of effective population size, $N_e$

Population size has a fundamental effect on random variation in the population mean from generation to generation. Such random or stochastic effects can be conveniently modeled using a single number called effective population size,  $N_e$  (Wright 1931, Barrowclough & Lande 1987, Charlesworth 2009). In the simplest kind of model, with no selection and equal family sizes, this number simply represents the number of parents. In more complicated models, effective population size can incorporate the sex ratio of breeding parents and variation in family size (Appendix 3). In all cases, however, the basic idea is that a single number,  $N_e$ , can account for stochastic variation in genetic properties that arises each generation from finite sampling of parents.

An evolving lineage is more than a single panmictic population with a particular  $N_e$ . In particular, we want to consider the case of a population subdivided into a set of demes of various effective sizes and exchanging migrants. As we follow this metapopulation through evolutionary time, some demes go extinct and their territory is colonized by other, more successful demes. Our intuition tells us that the total  $N_e$  for such a lineage composed of 10 demes, each with an  $N_e$  of 100, must be more than 100, perhaps more than 1000. For a particular model of migration between demes this intuition is known to be correct. Wright (1939; see Whitlock & Barton for ref; see also paper by Wright in 1942 with same title in Bull Math Stat) showed for the island model (... explicate ...) that the effective size of the total population is

$$N_e = \frac{n\bar{N}}{(1 - F_{ST})}, \quad (7.0)$$

where  $n$  is the number of demes,  $\bar{N}$  is their average effective size, and  $F_{ST}$  and  $(1 - F_{ST})$  are, respectively, the among- and within-deme components of variance in gene frequency. In our thought experiment,  $N_e$  could be as small as 1000 if all of the variation was within demes ( $F_{ST} = 0$ ), or large as 100,000 if nearly all of the genetic variation was among demes ( $F_{ST} = 0.99$ ). When  $F_{ST}$  is this large, migration is slight, and the demes behave as nearly independent entities. Under more general assumptions, Whitlock & Barton (1997) have shown that

$$N_e = \frac{n\bar{N}}{\{1 + Var(\mathcal{G})\}[\sum_i \mathcal{G}_i^2(1 - F_{STi})]}, \quad (7.1)$$

where  $\mathcal{G}_i$  is the eventual contribution of the  $i$ th deme to the whole population,  $Var(\mathcal{G})$  is the variance in those eventual contributions (which are scaled so that  $\sum_i \mathcal{G}_i = n$ ), and  $F_{STi}$  is the  $F_{ST}$  of the  $i$ th deme (Whitlock & Barton 1997). In this expression  $\bar{N}$  is an average weighted by  $\mathcal{G}_i^2$ . We see that (7.1) reduces to (7.0), if the eventual contributions of all demes are equal ( $Var(\mathcal{G}) = 0$ ), but in general  $Var(\mathcal{G})$  reduces total  $N_e$  below the simple expectation of  $n\bar{N}$ .

{end this section by applying the theory just developed to the case of  $N_e$  for evolution in an entire species = *T. elegans* ... lower bound from microsat data = Manier et al ... knock the estimate up one, two or three orders of magnitude to account for  $N_e$  across the geographic range ... knock the estimate down one or two orders of magnitude to account for temporal variation in  $N_e$  = geometric mean ... arrive at a final rough estimate, which will be used later in a MIPoD analysis; then bridge to the next section}

### 7.1 Estimates of $N_e$

{see table below from Charlesworth 2009 = long term, species wide estimates; find a review and incorporate results for short-term estimates from single local populations or arrays of populations; refs from E&A 2007 = Barrowclough 1980; Begon et al. 1980; Husband and Barrett 1992; Jorde and Ryman 1996; Storz et al. 2001; Turner et al. 2002; develop the distinction between these two kinds of estimates early on in this section!}

Table 1 | Effective population size ( $N_e$ ) estimates from DNA sequence diversities

Species	$N_e$	Genes used	Refs
<i>Species with direct mutation rate estimates</i>			
Humans	10,400	50 nuclear sequences	145
<i>Drosophila melanogaster</i> (African populations)	1,150,000	252 nuclear genes	108
<i>Caenorhabditis elegans</i> (self-fertilizing hermaphrodite)	80,000	6 nuclear genes	41
<i>Escherichia coli</i>	25,000,000	410 genes	146
<i>Species with indirect mutation rate estimates</i>			
Bonobo	12,300	50 nuclear sequences	145
Chimpanzee	21,300	50 nuclear sequences	145
Gorilla	25,200	50 nuclear sequences	145
Gray whale	34,410	9 nuclear gene introns	147
<i>Caenorhabditis remanei</i> (separate sexes)	1,600,000	6 nuclear genes	43
<i>Plasmodium falciparum</i>	210,000–300,000	204 nuclear genes	148

For data from genes, synonymous site diversity for nuclear genes was used as the basis for the calculation, unless otherwise stated.

{End this section by applying the lessons of (7.1) to the case of Eagle Lk *T. elegans* metapopulation, bringing in the consequences of source-sink pop dynamics etc; Manier et al 2005 and ms on stochastic demography in the same system}

### 7.2 Sampling from a normal distribution of breeding values

If we take repeated samples of size  $N$  from a normal distribution with variance  $V$ , the variance among the means of those samples will be  $V/N$ . This result reinforces our intuition that differences due to sampling will be appreciable with small samples, but inconsequential if the samples are large. We can apply this result to model the consequences of drift for replicate populations that obey the same rules of inheritance and sampling. We imagine a base population in which breeding values are normally distributed with a mean  $\bar{x} = \bar{z} = 0$  and a variance  $G$ . Suppose we establish replicate populations, each of size  $N_e$ , from this base population. The variance among those replicates in mean breeding value will be

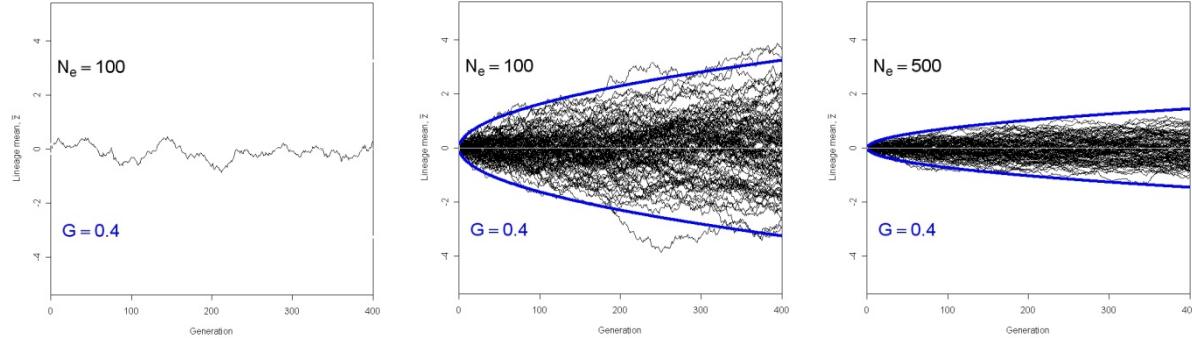
$$Var(\bar{z}_l) = \frac{1}{N_e} G \quad (7.00)$$

and the overall mean will be zero (Lande 1976).

### 7.3 Projecting the distribution of breeding values into the future

Now allow each of the replicate populations to constitute a lineage. In each lineage  $G$  remains the same and the same process of sampling occurs each generation. If we follow a single lineage through time, its phenotypic mean each generation,  $\bar{z}_t$ , deviates from the mean in the preceding generation,  $\bar{z}_{t-1}$ , by an

amount that can be specified by taking a random draw from normal distribution of breeding values with zero mean and variance  $G/N_e$ . Following the mean through time, we see a random walk (Fig. 7.0a). Each



**Figure 7.0** Genetic drift causes the phenotypic mean of a lineage to undergo a random walk. Simulations of genetic drift using 7.01 and 7.02. The lineage mean,  $\bar{z}$ , is shown in units of within-population phenotypic standard deviation,  $\sqrt{P}$ . (a) The random walk of a single lineage of small effective size. The random walks of 100 replicates lineages are shown for small (b) and moderate sized (c) populations. The theoretical 99% confidence limits for the mean of lineage means are shown in blue.

replicate lineage undergoes an independent random walk that obeys the same rules. We now wish to consider the statistical properties of the entire ensemble of replicate lineages at some generation,  $t$ , in the future. We began in generation 1 with an among-replicate distribution of means that was normally distributed with a mean of 0 and a variance of  $G/N_e$  (7.00). The distribution of means in generation 2 is the product of that distribution and normal distribution with the same properties, a so-called convolution of normal distributions, which is itself a normal distribution. By extending this logic, we can conclude that the among-replicate distribution of means at any future generation  $t$ , is a normal distribution

$$\Phi(\bar{z}_t) = \frac{\exp[-\frac{1}{2}(\bar{z}_t - \bar{z}_0)^2 / D(t)]}{\sqrt{2\pi D(t)}}, \quad (7.01)$$

with a mean of zero,  $\bar{z}_0 = 0$ , and a variance given by

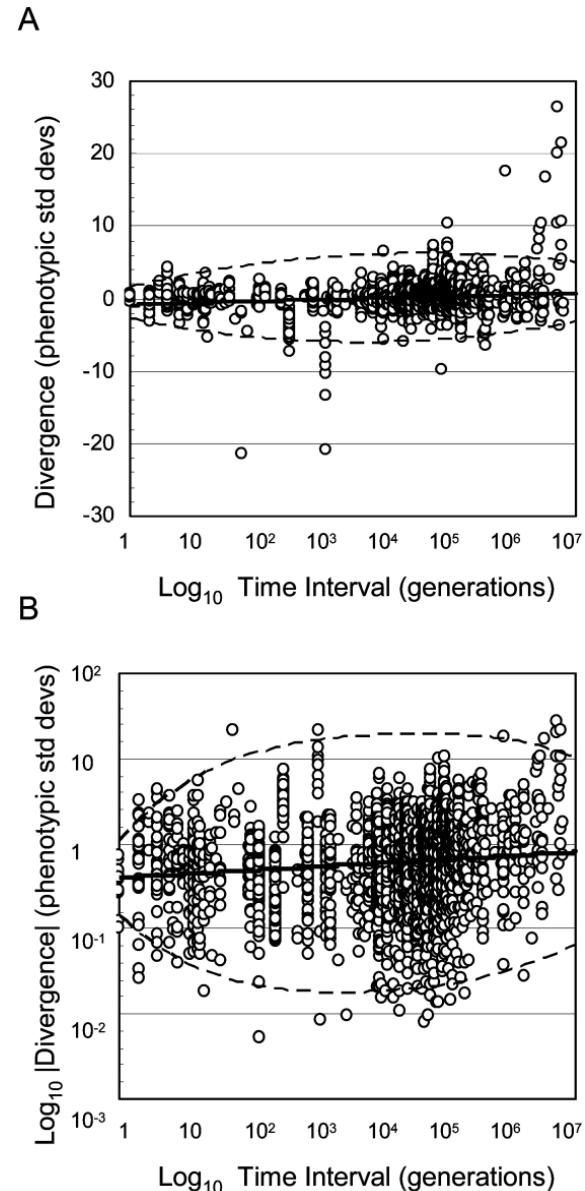
$$D(t) = \frac{1}{N_e} G t, \quad (7.02)$$

where  $D(t)$  stands for the dispersion (variance) of replicate means at generation  $t$  (Lande 1976). In other words, the variance among replicates increases linearly with time. This linear property means that the 99% confidence limits for the mean of means are ever-expanding, so that their graph forms a conic section (Fig. 7.0b, c) ... normal distributions of replicate means with increasing variance (Fig. 7.1, based on figs. in Lande 1976 and Estes & Arnold 2007).

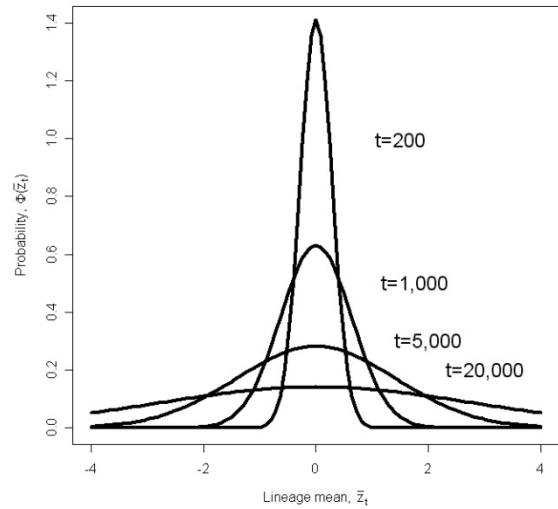
The results shown in Figs. 7.0 and 7.1 are obviously simplifications since we have ignored the effects of sampling on  $G$ , a problem to which we will return in Chapter y. We note, however, that (7.02)

might be a reasonable approximation if  $N_e$  is large enough so that the sampling effects on  $G$  are relatively small.

**Figure 7.1.** Distributions of lineage means under neutrality. Each curve represents the distribution of lineage means after some number of elapsed generations,  $t$ . Genetic variance = 0.4 and  $N_e = 1000$ .



**Figure 1:** A, Plot of divergence as a function of time interval. Divergence is measured as the difference between the average trait values of an ancestral and a descendant population, expressed in units of pooled phenotypic standard deviation. The slope of the fitted regression line is 0.2407 (SE = 0.0268),  $P < .0001$ . B, Plot of  $\log_{10}$  absolute divergence on  $\log_{10}$  time interval. The slope of the fitted regression line is 0.0437 (SE = 0.0070),  $P < .0001$ . In both plots, time interval is measured in number of generations on a  $\log_{10}$  scale. The broken lines show the boundaries of the 99% confidence ellipse for the data. The data ( $n = 2,639$ ) are from Gingerich (2001). The fitted slopes are equivalent to 0.84 phenotypic standard deviations per million generations on a raw scale.



We have also used a single number to represent population size throughout time in all replicate lineages, another obvious simplification. If we model the effects of temporal fluctuation in population size, it turns out that geometric mean population size provides a good overall approximation of sampling effects (Crow & Kimura 1970 or appropriate S. Wright ref). Bottlenecks have a disproportionately large effect that is not captured by the arithmetic mean but is captured by the geometric mean. Consequently, a single number, call it  $N_e$ , can be used to represent population size even when size fluctuates.

{End this section with a historical note on Brownian motion, ref general idea of stochastic processes, and general parameterization of BM in particular}

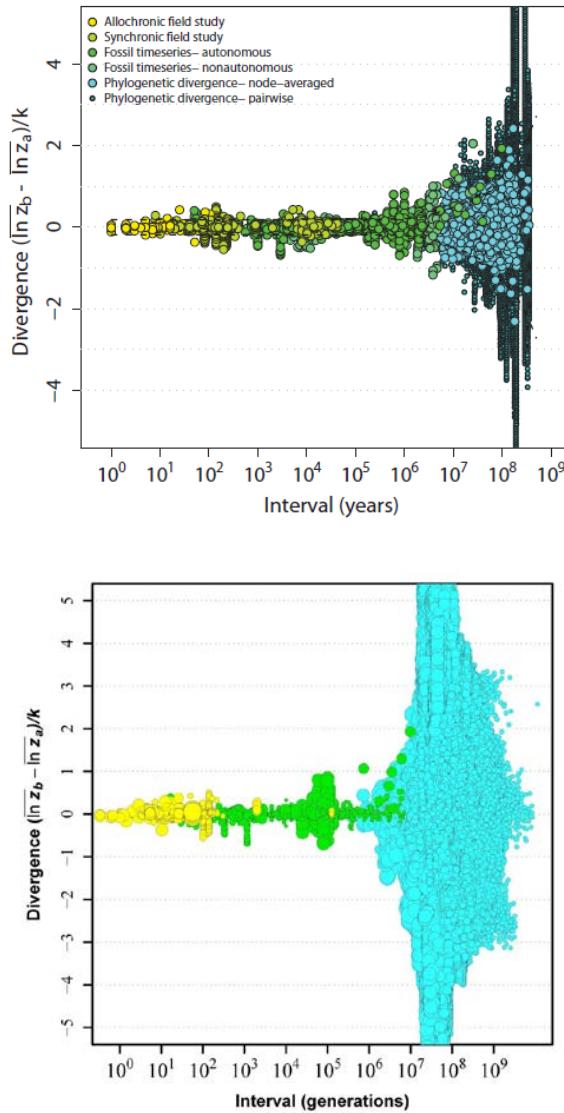
#### 7.4 Tests for neutral evolution with large data sets

... tests on long time scales ... Lynch 1990 ... test for neutrality with Gingerich data (Estes & Arnold 2007)

... Fig. 7.2 (based on neutrality figure in Estes & Arnold 2007 spreadsheet).

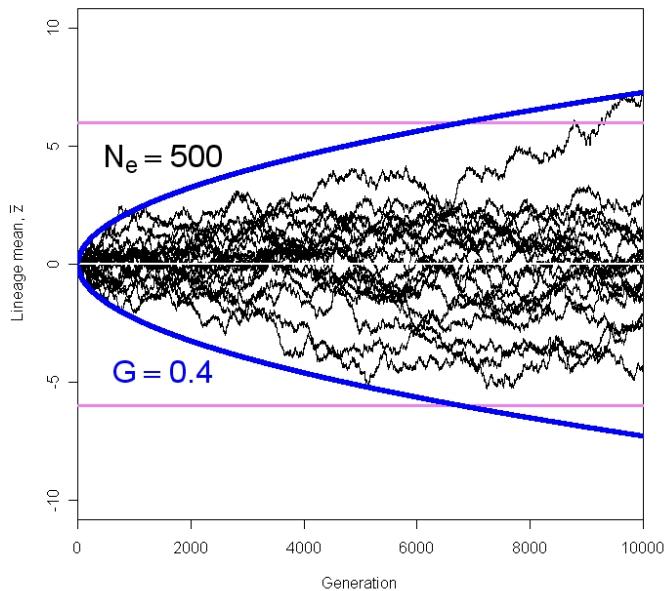
Fig. 7.x The relationship between evolutionary divergence and elapsed time. Divergence is measured as the difference between log-transformed size measurements in two populations, standardized by the dimensionality of the measurements. The interval on the x-axis is the time separating the two samples. (a) Divergence as a function of elapsed time measured in years.

(b) Divergence as a function of elapsed time measured in generations. From Uyeda et al 2011



The problem that genetic drift faces as a model of evolution can be seen as soon as we graph the process on long time scales. For example in Figure 7.5 we examine the consequences of drift on a moderately long time scale, using the same parameters that we employed in Fig. 7.0c on a short time scale. Although the process appeared restrained on a time scale of hundreds of generations, when we expand the time scale to thousands of generations, a substantial proportion of lineages diverge beyond the limits observed in large data sets (Fig. 7.5). On a time scale of millions of generations, this problem is exacerbated. Even populations with relatively large population sizes ( $N_e = \text{xxxxxx}$ ) and minuscule genetic variances ( $G = \text{xxxx}$ ) can drift beyond observed limits after 1-100 million generations (Estes & Arnold 2007).

**Figure 7.5.** A simulation of drift in 20 replicate populations for 10,000 generations based on 7.01 and 7.02. The blue lines show the theoretical 99% confidence limits for the mean of lineage means. The violet lines show the limits of phenotypic divergence that are observed on time scales less than 1 million generations (Estes & Arnold 2007, Uyeda et al. 2011), namely, less than  $\pm 6$  within-population phenotypic standard deviations,  $\sqrt{P}$ .



## 7.5 Drift on a phylogeny

Another simplification in (7.02) is that we imagined that the replicate lineages diverge from a single common ancestor at generation 0. In other words, the replicates evolved on a socalled star phylogeny in which all the branch lengths  $t$  generations. In applying neutral theory to evolving populations we will often want to account for a more complex phylogeny.

Independent contrasts represent one way to take account of phylogeny (Felsenstein 1985). To apply this method, one converts data at the tips into contrasts (differences between means) that can be assumed to be independent if the stochastic process of trait evolution on the tree is Brownian motion (e.g., drift, as characterized above). For the tree shown in Fig. 7.3, the independent contrasts would be  $(\bar{z}_b - \bar{z}_c)$  and  $(\bar{z}_a - \bar{z}_{bc})$ , where  $\bar{z}_{bc}$  is the trait mean of the common ancestor of  $b$  and  $c$ , which we will call  $bc$ . In data analysis, under the assumption of Brownian motion, these two contrasts are treated as independent observations of evolutionary outcome and scaled so that they have same expected trait variance. An alternative approach avoids the process assumption of independent contrasts as well as the need to reconstruct the means of common ancestors. In this alternative approach, we reconstruct the trait variance-covariance matrix for taxa evolving on the tree. This matrix could be based on Brownian motion or any other well-characterized stochastic process. In other words, we derive expressions for the off-diagonal elements of this matrix rather than transforming to contrasts so that they are zero (independent). A general solution to those elements has been derived by Hansen & Martins (1996).

Hansen & Martins (1996) have shown that the expected covariance of the trait means of two taxa evolving according to a stochastic process on a tree is equal to the expected trait variance of their most recent common ancestor. Suppose that ancestor is  $t$  generations removed from the ancestor of the tree. Then, if the stochastic process is drift, the trait variance of the common ancestor is the variance among replicates of that common ancestor evolving by drift for  $t$  generations is given by (7.02). We can use this

result to calculate the trait variance-covariance matrix for all of the taxa represented on the tips of the branches of a phylogeny.

To get an expression for that variance-covariance matrix, which we will call  $\mathbf{A}$ , we proceed in two steps. First, we represent the phylogeny as a matrix that gives the elapsed time from the tree ancestor to the tip taxa and to all of their common ancestors. Suppose there are  $r$  tip taxa. We can represent the phylogeny of those taxa with an  $r \times r$  matrix  $\mathbf{T}$ , in which the off-diagonal  $ij$ th element is the time (in generations) from the root of the tree to the most recent common ancestor of taxon  $i$  and taxon  $j$ , and the diagonal elements are the time from the root to extant taxon  $i$ .  $\mathbf{T}$  is a matrix of shared ancestry in the sense that the elements represent the number of elapsed generations in which pairs of taxa experienced shared ancestry (Martins 1995, Hansen & Martins 1996). For example, the phylogeny of three taxa ( $a, b, c$ ) in Fig. 7.3 can be represented by

$$\mathbf{T} = \begin{bmatrix} T_{aa} & T_{ab} & T_{ac} \\ T_{ba} & T_{bb} & T_{bc} \\ T_{ca} & T_{cb} & T_{cc} \end{bmatrix} = \begin{bmatrix} 1000 & 200 & 200 \\ 200 & 1000 & 800 \\ 200 & 800 & 1000 \end{bmatrix} \quad (7.04)$$

In the second step, we describe the outcome of drift on a phylogeny from generation 0 to present time, at the tips of the branches, by producing an  $r \times r$  matrix analog of (7.03), viz.

$$\mathbf{A} = \frac{1}{N_e} G \mathbf{T}, \quad (7.05)$$

where  $G$  represents the average additive genetic variance for the trait across time and across populations,  $N_e$  is a similar average for effective population size, and  $\mathbf{T}$  is the  $r \times r$  matrix of shared ancestry (Hansen & Martins 1996, Hohenlohe & Arnold 2008). The off-diagonal,  $ij$ th element of  $\mathbf{A}$  is the expected covariance in mean trait values between two taxa,  $i$  and  $j$ , with a shared coancestry equal to the  $ij$ th element of  $\mathbf{T}$ . Similarly, the diagonal,  $ii$ th element of  $\mathbf{A}$  is the expected variances in mean trait values after an elapsed time given by the  $ii$ th element of  $\mathbf{T}$ . For example, the  $\mathbf{A}$ -matrix for drift of a trait on the three taxon phylogeny shown in Fig. 7.3, with  $N_e=100$ , is

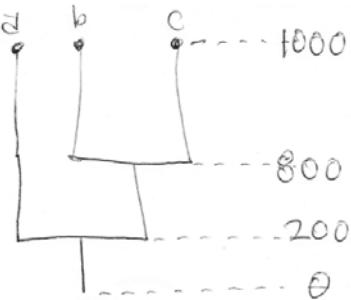
$$\mathbf{A} = \begin{bmatrix} (1000/100)G & (200/100)G & (200/100)G \\ (200/100)G & (1000/100)G & (800/100)G \\ (200/100)G & (800/100)G & (1000/100)G \end{bmatrix} = \begin{bmatrix} 10G & 2G & 2G \\ 2G & 10G & 8G \\ 2G & 8G & 10G \end{bmatrix} \quad (7.06)$$

In other words, if a series of replicate populations of taxon  $a$  evolved by drift on the phylogeny, the expected variance among their trait means would be ten times the genetic variance of the trait,  $10G$ . If a series of replicate pairs of taxa  $a$  and  $b$  evolved by drift on the phylogeny, the expected covariance of their trait means would be  $2G$ .

With replicate taxa that have evolved to the tips of a phylogeny given by  $\mathbf{T}$ , the distribution of traits means,  $\bar{z}$ , is normal,

$$\Phi(\bar{z}) = \frac{\exp[-\frac{1}{2}(\bar{z} - \bar{z}_0)^2 A^{-1}]}{\sqrt{(2\pi)^r |A|}}, \quad (7.07)$$

with a mean the same as in the ancestral population,  $\bar{z}_0$ , and a variance-covariance matrix given by  $A$ .



## 7.6 Tests of neutral evolution on trees

... large literature from comparative studies {the problem we confront here is that nearly all of these studies adopt a model of neutral evolution (Brownian motion) without testing it; might be able to use published values of the fitted drift variance parameter to show that they are unrealistic; see Harmon et al. 2010} ... {the issue of rate parameter varying across branches; perhaps a summary table for this section}

## 7.7 Mutation-migration-drift balance in a metapopulation.

The complications of phylogeny disappear when we consider a metapopulation subdivided into a series of populations that exchange migrants each generation. The ancestry of the populations becomes immaterial because we assume that the input of variation from mutation and the tendency to homogenize population differences by migration have reached a steady state. Once the metapopulation has equilibrated, how will the total genetic variation in a selectively neutral trait in this metapopulation be apportioned among its constituent populations? We can get a useful answer this question, by making a series of simplifying assumptions. Let us assume that there are  $n$  populations each of effective size  $N$  and that the rate of migration between populations (the fraction of individuals that move from one population to another population each generation),  $m$ , is the same in both directions between all pairs of populations. We will also assume our standard additive genetic model for inheritance and that mutation contributes a constant amount,  $U$ , to the genetic variance in each population each generation. Under these conditions, Lande (1992; reconcile the results that follow with Whitlock 1999 Genet Res which are general for migration and case in terms of  $F_{ST}$ ) has shown that the genetic variance within ( $G$ ) and among populations ( $G_a$ ) equilibrates so that

$$\hat{G} = 2nNU \quad (7.08)$$

and

$$\hat{G}_a = \frac{n-1}{m}U \quad (7.09)$$

Twice the genetic variance within populations is converted into genetic variance among populations by the process of random drift (Wright 1951, 1969). Consequently, the total genetic variance in the metapopulation is  $G_a + 2G$  rather than  $G_a + G$ . Several important conclusions can be drawn from these equilibrium results. First, the migration rate,  $m$ , does not affect the level of genetic variance within populations. This results contradicts our intuition that migration should enhance variation within populations, but we have assumed that equal numbers of migrants enter and leave each population. Second, the level of genetic variation maintained among populations is enhanced by increasing the number of populations, eroded by increasing migration rate but is not affected by population size,  $N$ . The rates of approach to these equilibria (not shown) are negative exponential functions, so that the approach is rapid at first, following by a slow, asymptotic period. With a large number of populations, that final approach takes a few times  $n/2m$  or  $2nN$  generations, whichever is larger (Lande 1992). So, for 20 populations of effective size 500, exchanging  $mN=2$  migrants per generation ( $m=1/250$ ), several thousand generations might be required for the system to equilibrate in mutation-migration-drift balance {would be easy to tailor these numbers to fit the elegans metapopulation of Manier et al 2007!}.

It is useful to express the among-population genetic variance as a fraction of the total,

$$Q_{ST} = G_a / (G_a + 2G), \quad (7.10)$$

a fraction comparable to Wright's  $F_{ST}$ , a measure of population differentiation in gene frequencies (Wright 1969, Lande 1992, Spitzke 1993). ... (Fig. 7.4) = Multiple panel figure showing within- and among-population distributions of genetic values as a function of  $Qst$ . When the two parts of variation in a quantitative trait are estimated by a two-level analysis of variance,  $G_a$  and  $G$  are, respectively, the among- and within-population components of genetic variance.

{add material from more recent paper by Whitlock}

### 7.8 Tests for neutral evolution in metapopulations

.... tests for neutral evolution of garter snake coloration and vertebral numbers (Manier et al. 2007) – report in text just the results of the  $Fst/Qst$  comparison without going into the 3-level testing, don't attempt to produce a table, again too complicated; present and discuss Fig. 7.5 = Distribution of  $Fst/Qst$  results, produced using summary table in review(s) by Merila et al. = just  $Fst/Qst$  ... review of  $Fst/Qst$  studies; need to find most recent review, see Manier et al. 2007 for ref to review by Merila et al ;

## Chapter 11: Evolution of a Single Trait on a Stationary Adaptive Landscape

© Stevan J. Arnold 2016

**Overview.-** The adaptive landscape is a powerful conceptual tool for modeling long-term responses to selection. In the simplest case, we imagine a set of replicate populations that descend all at once from a common ancestor and thereafter evolve on identical adaptive landscapes. As a general rule, variation among replicate means is inversely proportional to effective population size. The smaller the size of the replicates, the more variable their response to selection. A  $\cap$ -shaped adaptive landscape results in a stable equilibrium distribution of replicate means. Stabilizing selection tends to pull means towards an optimum, but that tendency is balanced by drift away from the optimum. Tests with a large data set indicate that this Ornstein-Uhlenbeck (OU) model is undoubtedly an element in the general explanation for stasis, but other factors are also in play.

### 11.0 Tendency to evolve uphill on the adaptive landscape

$$\Delta\bar{z} = G\beta = G \frac{\partial \ln \bar{W}}{\partial \bar{z}}$$

... Fig. 11.0 ...

**11.1 Stochastic dynamics and equilibrium on a landscape with a single, stationary adaptive peak**  
 Gaussian adaptive landscape ... linear restoring force of selection towards the optimum proportional to the distance from the optimum Fig. 11.1 = Rubberband cartoon of the OU process \*\*before this

figure need to specify that restoring force as  $\frac{(\bar{z} - \theta)}{\omega + P}$  and identify it as the force of directional

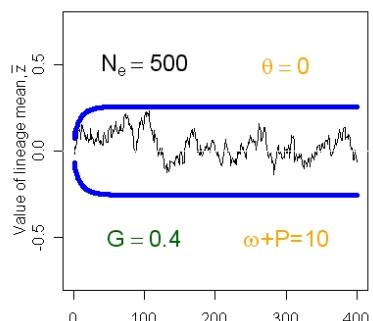
selection back towards the optimum,  $\theta$ , is  $\beta = \frac{(\bar{z} - \theta)}{\omega + P}$  \*\*... replicate populations of size  $N_e$  ... Ornstein-

Uhlenbeck (OU) process (Ornstein & Uhlenbeck 1930) ... each generation the trait mean drifts by an amount equivalent to a draw from a normal distribution with mean zero and variance  $G/N_e$  and is pulled back towards the optimum with a force,  $\beta$ , that is proportional to the deviation from the optimum. In other words,

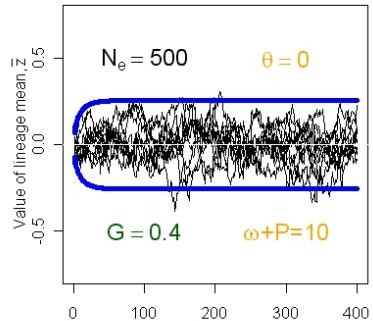
$$\bar{z}(t+1) = \bar{z}(t) + \frac{(\bar{z}(t) - \theta)}{\omega + P} G + N(0, G/N_e)$$

... let the position of the optimum be zero ... after  $t$  elapsed generations, the expected value of the mean of lineage means is zero with lineage means normally-distributed about that expectation with a variance of

$$Var(\bar{z}_t) = \frac{\omega + P}{2N_e} \left\{ 1 - \exp \left[ -2 \left( \frac{G}{\omega + P} \right) t \right] \right\}.$$



An example of a lineage drifting about its optimum is shown in Fig. 11.3a. Each time the trait mean drifts away from the optimum, it is



pulled back, so ,as a long term average, the mean resides at the optimum. An ensemble of 10 replicate populations, buzzing about identical optima are shown in Fig. 11.3b. The outer limits of their paths are accurately described by the 99% confidence limits (shown in blue).

**Fig. 11.3.** Simulations of replicate populations evolving about a stationary intermediate optimum. (a) A single lineage mean in a population of finite size evolving according to an O-U process. (b) The lineage means of 10 replicate populations evolving according to the same process.

At equilibrium, trait means are normally distributed with a mean of means at the optimum,  $\theta$ , and a variance of lineage means given by

$$\text{Var}(\bar{z})_{\infty} = \frac{\omega + P}{2N_e},$$

so that the equilibrium distribution of means is given by

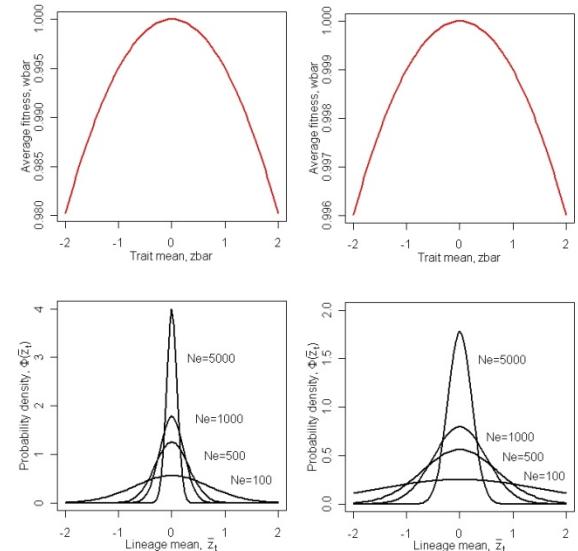
$$\Phi(\bar{z})_{\infty} = \frac{1}{\sqrt{2\pi\text{Var}(\bar{z})_{\infty}}} \exp\left\{-\frac{(\bar{z}-\theta)^2}{2\text{Var}(\bar{z})_{\infty}}\right\}.$$

Lande (1976). In other words, variation in lineage means is greater if stabilizing selection is weak (large  $\omega + P$ ) and populations effective size is small (Fig. 11.2). Smaller populations can drift further from the optimum, especially if the restraining force of selection is weak.

**Figure 11.2** Equilibrium distributions of lineage means under weak stabilizing selection. The upper panels show the adaptive landscape (note change in vertical scale from left to right). The lower panels show equilibrium distributions of lineage means,  $\Phi(\bar{z}_{\infty})$ , as a function of effective population size,  $N_e$ . (a) Weak stabilizing selection,  $\omega + P = 100$ . (b) Very weak stabilizing selection,  $\omega + P = 500$ .

On a more quantitative level, these equations show that stabilizing selection must be extremely weak to produce appreciable variation in lineage means at equilibrium (Fig. 11.2). For example, even a restraining force as weak at  $\omega + P = 500$ , dispersion of lineage means is small, so that even with an effective size of 500, the standard deviation of lineage means is less than 1 (compare with the drift distributions in Fig. 7.1).

... early-burst model (Blomberg et al. 2003), also known as the ACDC model ... variance among lineages decreases through time to account for bursts of diversification early in adaptive radiations ... at rate  $r$  ( $r < 0$ ) ...



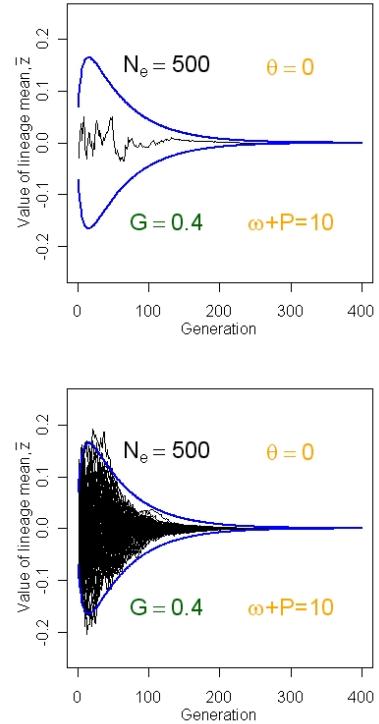
$$\bar{z}(t+1) = \bar{z}(t) + \frac{(\bar{z}(t) - \theta)}{\omega + P} G + N(0, \text{var})$$

where  $\text{var} = (G/N_e)\exp\{rt\}$  and  $r < 0$  is a parameter describing the diminution of drift variance with time (Harmon et al. 2010). The expected value of the lineage means is  $\theta$  with normally-distributed variation about that value so that variance at generation  $t$  is

$$\text{Var}(\bar{z}_t) \approx \frac{\omega + P}{2N_e} \left\{ 1 - \exp \left[ -2 \left( \frac{G}{\omega + P} \right) t \right] \right\} [\exp\{0.7rt\}] .$$

From this equation we see that the variance among replicate lineages decays exponentially. This equation differs from results given by Blomberg et al. (2003) and Harmon et al. (2010) by incorporating realistic parameters for drift and stabilizing selection. Taking the same parameters for population size, selection, and inheritance as in the last example, Fig. 11.8 portrays an example of results from the early-burst model. As expected, the model produces about as much among-lineage variation as the simple OU model from generations 0 to about 20, but after that variance rapidly decays so that after just a few hundred generations, all replicate lineage mean reside almost exactly at the optimum. These features persist so long as  $r$  is negative and different from zero, although as  $r$  approaches zero, (xEB var = last expression) converges on (xOU var).

In comparing the pattern in Fig. 11.8 with the pattern of data on long time scales (Fig. 7.x), we must remember that the onset of radiations are not synchronized to occur at the beginning of time intervals, as they are in the simulations. Instead, according to the model, we might expect bursts of diversification, followed by exponential decay, to periodically occur in the data. We must also realize that the early-burst model produces a radiation of modest proportions (i.e., far less than  $\pm 6\sqrt{P}$ ) followed by collapse towards virtually no variation about the optimum. Thus, the early-burst model, unless supplemented by peak movement, cannot produce the full width of the long-term band apparent in Fig. 7.x and totally fails to produce the surge in diversification beyond  $\pm 6\sqrt{P}$  that is apparent after long time intervals. Not surprisingly, Harmon et al. (2010) found that the early-burst model produced worse fits to data than Brownian movement and an OU process.



**Fig. 11.8.** Simulations of lineages evolving according to the early-burst model of Blomberg et al. (2003). Conventions as in Fig. 11.2. The parameter describing diminution of drift variance,  $r$ , is set to -0.05. (a) A simulation of one evolving lineage. (b) 100 replicate lineages.

## 11.2 Exploration of a stationary landscape by drift

... interaction of drift and selection ... modeling the change in the distribution with a diffusion equation ... the change in the distribution of replicate population means through time is given by

$$\frac{\partial \Phi}{\partial t} = -\frac{\partial}{\partial \bar{z}_t} \left\{ -\frac{G}{(\omega + P)} \bar{z}_t \right\} + \frac{\partial^2}{\partial \bar{z}_t^2} \left\{ \frac{G}{N_e} \right\}$$

where the terms in brackets are, respectively, the mean and variance of the transition probability (per generation) density function for  $\Phi(\bar{z}_t)$  (Lande 1976).

... at equilibrium

$$\Phi(\bar{z})_\infty \propto \bar{W}^{2N_e}$$

... expected first passage time

Lande (1976).

{refer back to Fig. showing sample univariate paths of  $\bar{z}$  on a univariate Gaussian AL}.

### 11.3 Attempts to explain time series and adaptive radiations with a single stationary peak

... Estes & Arnold 2007 ... Gingerich data and a single peak (Fig. 11.4 shows tests from Estes & Arnold 2007 spreadsheet) ... tree-based data (Hansen 1997; Butler & King 2004, OUCH; Harmon et al 2010; the necessity of assuming very small  $N_e$  and/or extremely weak stabilizing selection;

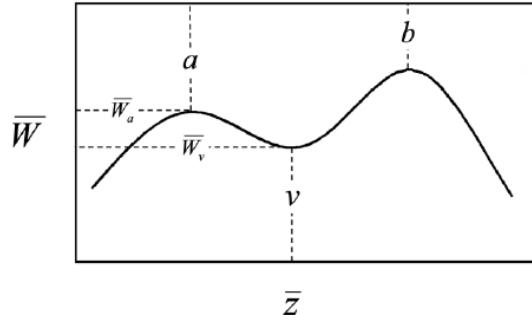
### 11.4 Shifting of the trait mean between two adaptive peaks

... Fig. 11.5 from Lande 1985 or 6

$$T \cong \frac{2\pi(\omega + P)}{G} \left( \frac{\bar{W}_a}{\bar{W}_v} \right)^{2N_e}$$

$$E[\bar{z}_t] = d(1-p)$$

$$Var(\bar{z}_t) = p\bar{z}_t^2 + (1-p)(d - \bar{z}_t)^2 + \frac{\omega + P}{2N_e}$$



... the first term on the right ... the second term ... the third term ... (Lande 1985, 1986).

... Estes & Arnold (2007) tests with the Gingerich data  
... need to invoke unrealistically small  $N_e$

**Figure 5:** Adaptive landscape in the peak shift model, which has two adaptive peaks,  $a$  and  $b$ . Mean fitness  $\bar{W}$  is a function of the average phenotype in the population,  $\bar{z}$ . Two critical parameters in the peak shift model that affect the probability that the population mean will shift from peak  $a$  (left) to peak  $b$  (right) are the height  $\bar{W}_a$  of the first adaptive peak and the height  $\bar{W}_v$  of the valley  $v$  separating the first adaptive peak from the second. The distance between the two peaks (between points  $a$  and  $b$ ) is  $d$ .

### 11.5 Geographic variation

Geographic variation refers to among-population variation within a species. Such variation is an almost ubiquitous feature of most species, although the extent of geographic variation varies from trait to trait (Mayr 1963). ... example plots from James' study of body size in NA birds ... an intermediate optimum that varies in space ... ... results for clines

## Chapter 12: Evolution of Multiple Traits on a Stationary Adaptive Landscape

© Stevan J. Arnold 2016

**Overview.-** The theoretical evolution of the multivariate mean has been most explored in the case of a hill-shaped adaptive landscape that has a consistent position, shape and orientation. Under certain conditions, the mean tends to evolve uphill on this landscape, towards the adaptive peak. ... evolution of two genetically correlated traits ... . Theory is especially well developed for systems in which a single male trait evolves in response to sexual selection exerted by a second, female trait (sexual preference). Under these conditions, the equilibrium of the bivariate mean may be a stable point, a line or an elliptical cycle. Unstable dynamics are also possible, but probably unlikely. Finite population size adds an elements of uncertainty to the outcome, changing stable points into clouds, for example. Predictions from these models are consistent with various features of populations with sexually-selected traits, but discriminating, quantitative tests have not yet been accomplished.

{in this introduction need to bridge from the generally pessimistic results of empirical tests with single traits, preceding chapter, to this chapter ... single peaks are an element in the explanation of adaptive radiations but not the whole picture ... }

### 12.0 The tendency to evolve uphill on the adaptive landscape

$$\Delta \bar{z} = G\beta = G \frac{\partial \ln \bar{W}}{\partial \bar{z}} \quad (12.00)$$

... Mahalanobis distance in phenotypic space ...  $[(\bar{z}_a - \bar{z}_b)^T P^{-1} (\bar{z}_a - \bar{z}_b)]^{\frac{1}{2}}$  ... *generalized phenotypic distance* between the means of two populations,  $a$  and  $b$ , ... accounts for covariances among traits and differences in variance ... for generalized *genetic* distance, replace  $P$  with  $G$  ... the square of generalized genetic distance is always positive and equivalent to the change in log mean fitness (Lande 1979),

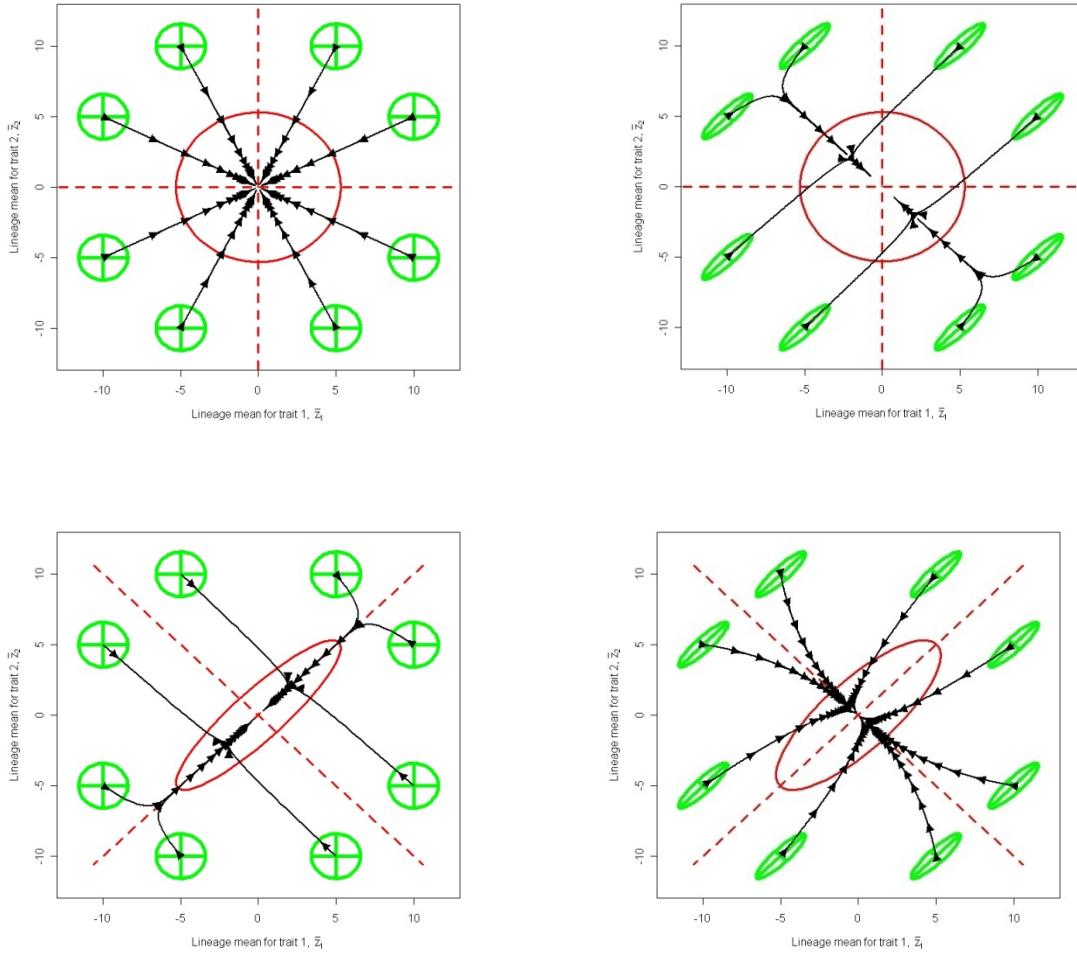
$$\Delta \ln \bar{W} \cong \Delta \bar{z}^T G^{-1} \Delta z \geq 0 \quad (12.01)$$

... consequently we can conclude that ... (Fig. 12.0 = figure from chapter in the Boake book).

### 12.1 The spaceship model

The case of two traits evolving on a stationary adaptive landscape illustrates many properties of the general multivariate case, so we will use it as a starting point. In the case of a single trait, the process under discussion consists of the boring march of the frequency distribution, disparaged by Luria et al. 1981 (*A View of Life*), but in the two trait case, the process is much more interesting. In particular, we shall consider the process when the bivariate means lies some considerable distance from the optimum. Under this circumstance we will see that the process consists of two distinct phases: an initial phase during which the mean evolves rapidly towards the optimum, and a second phase, often with a distinctly different trajectory, during which the mean slowly approaches the optimum. ... deterministic response ... . The per generation change in the trait mean when the AL is Gaussian is a special case of the general expression,  $G\beta$ ,

$$\Delta \bar{z} = G\beta = G(\omega + P)^{-1}[\theta - \bar{z}_t]$$



**Figure 12.1** Bivariate evolution on a Gaussian adaptive landscape, the spaceship model. The adaptive landscape is shown in red: the ellipse represents the equivalent 50% confidence region, with eigenvectors shown as dashed lines. Stabilizing selection of equal magnitude acts on each trait:  $(\omega + P)_{11} = (\omega + P)_{22} = 50$ . G-matrices (represented as 95% confidence ellipses) are shown in green. In each case, the genetic variance of both traits is 0.4. Evolutionary trajectories of the bivariate mean are shown in black, with arrows at regular intervals of elapsed generations. (a)  $r_g = 0, r_s = 0$ . Each evolutionary trajectory is 500 generations in duration, with arrowheads every 50 generations. (b)  $r_g = 0.9, r_s = 0$ . Each evolutionary trajectory is 1500 generations in duration, with arrowheads every 300 generations. (c)  $r_g = 0, r_s = 0.9$ . Each evolutionary trajectory is 500 generation in duration, with arrowheads every 50 generations. (d)  $r_g = 0.9, r_s = 0.8$ . Each evolutionary trajectory is 700 generations in duration, with arrowheads every 50 generations.

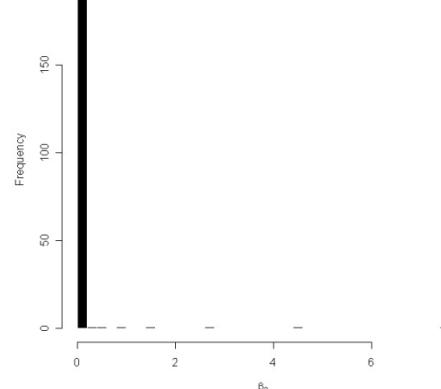
It will be useful to remember the two trait version of this last equation as the ‘spaceship model’. Think of a bivariate mean located some considerable distance from a single, stationary adaptive peak. The mean is surrounded by a cloud of genetic values that can be visualized as a green ellipse, our spaceship. In a thought experiment, we keep the overall configuration of the cloud constant as the bivariate mean responds to selection exerted by the distant peak and moves toward it. The particular trajectory that the spaceship takes on its journey is a function of its starting position, as well as the configurations of the genetic cloud and the AL (Fig. 12.1)

Fig. 12.1 illustrates three important properties of the lineage mean’s response to an distant adaptive peak. First, responses are greater when the mean is further from the adaptive peak. This greater response reflects the fact that directional selection is stronger, the further away the mean is from the optimum. Second, genetic covariance and correlational selection can each cause evolutionary trajectories to be curved. Third, the evolving mean can over- or under-shoot the optimum so that the population is subjected to long periods of maladaptation (Fig. 12.1b, c). This phenomenon of maladaptation is likewise a consequence of genetic covariance and correlational selection. We will now examine these aspects of the spaceship’s trajectory and journey in greater detail.

We also see in Fig. 12.1 that the shapes of the G-matrix and the AL can dramatically affect the evolutionary trajectory of the mean during both the initial and secondary phases of approach to the optimum. In the absence of both genetic correlation and correlational selection, both the cloud of genetic values and the AL are circular in cross-section. Under this circumstance (Fig. 12.1a [currently named 12.1c]), the mean evolves directly towards the adaptive peak. Evolution is rapid at first and then gradually decelerates, until the mean ceases its evolution at the peak. If the AL is circular but the G-matrix is elongate, the rapid initial phase is in the direction of  $g_{max}$ , the long axis of the G-matrix (Fig. 12.1b [12.1b]). However, when the mean approaches  $g_{min}$ , the short axis of the G-matrix, the trajectory turns in that direction and evolution decelerates dramatically as the mean completes its slow approach to the optimum. The difference in initial and secondary phases makes intuitive sense: a rapid initial phase fueled by abundant genetic variance, then a slow secondary phase in the direction of least genetic variance. This intuition is expanded when we consider a third case, in which the G-matrix and the AL have reverse configurations (Fig. 12.1c [12.1a]). Now the genetic cloud is circular and impotent, and the elongate AL governs the trajectories of approach to the optimum. The rapid initial phase is along  $\omega_{min}$ , the short axis of the AL, and the slow, second phase is along  $\omega_{max}$ , the long axis of the AL. When genetic variance is the same in all directions, the shape of the AL dictates the difference in trajectories during the initial and secondary phases of evolution. Finally, in a revealing but unillustrated case, consider the outcome when the G-matrix and the AL are perfectly aligned ( $g_{max} = \omega_{max}$ ) and proportional (i.e., the corresponding eigenvalues of the two matrices differ by the same constant of proportionality). Under this circumstance of alignment and proportionality, the effects of the two process cancel and the trajectories of the mean assume the star-like pattern seen in Fig. 12.1a. In retrospect, we could have anticipated this result from (12.x): since  $cG = (\omega + P)$ , where  $c$  is a scalar constant,  $cG(\omega + P)^{-1}$  is  $cI$ , where  $I$  is the identity matrix. Because this special case requires a precise combination of parameters, a more general case is illustrated in Fig. 12.1d. Here, G and the AL are aligned but not exactly the same shape, with the result that both inheritance and selection affect the trajectories of the mean.

... {some quantitative statements about how long the first and second phases last} ...

A arbitrary feature of these illustrations of the spaceship model is that we began each trajectory with the

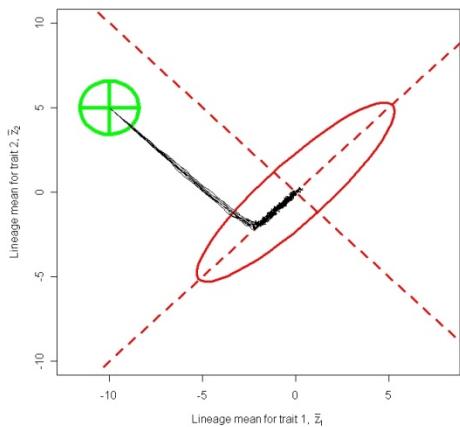


mean situated far from the optimum. How commonly does the trait mean reside 5-10 standard deviation away from the optimum? Under the Gaussian assumptions used in Fig. 12.1, the magnitude of the selection gradient,  $\beta$ , is directly proportional to the distance to the optimum and inversely proportional to the curvature of the AL,  $(\omega+P)$ . Consequently, we can produce a histogram of  $\beta$ -values for any of the trajectories shown in Fig. 12.1 and compare that histogram to the histogram of  $\beta$ -values estimated in natural populations (Fig. 2.4c). A representative histogram from a trajectory in Fig. 12.1 is shown in Fig. 12.2. That trajectory produces a few values of  $\beta$  that are much larger than those typically seen in natural populations and of course these values were produced during the first few generations of the first phase, when the mean was far from the optimum. The histogram also shows an overabundance of values very close to the optimum. In other words, the process of deterministic approach to an optimum from far away, as in Fig. 12.1, is not consistent with the picture of evolution implied by direct estimates of directional selection. ... this result highlights a problem with evolution along genetic lines of least resistance as an explanation for adaptive radiations ...

So far we have considered the deterministic evolution of the bivariate mean. Finite population size will cause stochasticity in the sampling of parents each generation and produce a distribution of trajectories. One such case is shown in Fig. 12.3. In general, as in this case, stochasticity is tempered by directional and stabilizing selection, so that the variation among lineage means is considerably less than in the neutral case.

**Figure 12.3** Stochastic response of the bivariate mean to a Gaussian adaptive landscape for 500 generations in 8 replicate lineages. No genetic correlation,  $r_s=0.9$  and  $N_e=500$ . Other conventions as in Figure 12.1 except that elapsed intervals are not indicated with arrowheads.

{a final paragraph for this section pointing out that even the 2-trait portrayal has implications for multiple trait evolution that go far beyond the specifics of sexual dimorphism ... general portrayal of the role of genetic covariance in evolution ... the real possibility of temporary but long lasting maladaptation ...}



## 12.2 The evolution of sexual dimorphism

Following Lande's (1980b) treatment of this problem, we will assume that multiple autosomal loci affect expression of the trait differently in the two sexes. Let  $z$  and  $y$  be normally distributed traits in males and females, respectively, of a dioecious species with within-sex genetic variances  $G_m$  and  $G_f$ , and a between-sex genetic covariance of  $B$ . The per generation evolution of the means,  $\bar{z}$  and  $\bar{y}$ , is a special case of (10.01),

$$\begin{bmatrix} \Delta\bar{z} \\ \Delta\bar{y} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} G_m & B \\ B & G_f \end{bmatrix} \begin{bmatrix} \beta_m \\ \beta_f \end{bmatrix} = \frac{1}{2} \begin{bmatrix} (G_m\beta_m + B\beta_f) \\ (B\beta_m + G_f\beta_f) \end{bmatrix}, \quad (12.02)$$

where the factor of  $\frac{1}{2}$  accounts for sex-limited expression;  $\beta_m$  and  $\beta_f$  are the selection gradients acting on the two sexes. { diagram, instead of equations for the two-sex adaptive landscape, illustrating the two thetas, two omegas + Ps } We assume that the two traits are under Gaussian viability selection towards intermediate optima ( $\theta_m$  and  $\theta_f$ ). Each generation viability selection is followed by a process of

assortative mating and frequency-dependent sexual selection. The total selection gradients in (12.02) are composed of two terms, representing the directional forces of viability and sexual selection.

$$\begin{bmatrix} \beta_m \\ \beta_f \end{bmatrix} = \begin{bmatrix} \frac{(\theta_m - \bar{z})}{(\omega_m + P_m)} + c_m \\ \frac{(\theta_f - \bar{y})}{(\omega_f + P_f)} + c_f \end{bmatrix} \quad (12.03)$$

If sexual selection does not affect mean fitness and only affects the distribution of offspring among the individuals of each sex, the adaptive landscape for the two sexes is defined solely by natural selection. For simplicity, we will assume that the forces of sexual selection acting on the two sexes,  $c_m$  and  $c_f$ , are constant from generation to generation.

Evolution ceases at an equilibrium given by

$$\bar{z} = \theta_m + (\omega_m + P_m)c_m \quad (12.04a)$$

$$\bar{y} = \theta_f + (\omega_f + P_f)c_f, \quad (12.04b)$$

where  $P_m$  and  $P_f$  are the within-sex phenotypic variances, while  $(\omega_m + P_m)$  and  $(\omega_f + P_f)$  are the ‘variances’ of the Gaussian adaptive landscape function in the  $\bar{z}$  and  $\bar{y}$  dimensions. Note that if no individual expresses both  $z$  and  $y$ , the phenotypic covariance between the two traits is undefined, and there is correlational selection between the two traits. Consequently, the corresponding adaptive landscape lacks tilt because  $\omega_{mf}=0$  (Fig. 12.1). We see from (12.04) that the two traits will equilibrate on their adaptive peaks if there is no sexual selection ( $c_m=0$  and  $c_f=0$ ) (Fig. 12.1a, b). On the other hand, if sexual selection acts only on males ( $c_f=0$ ), males will be pulled off their adaptive peak at equilibrium, but females will not (Fig. 12.1c,d).

To analyze the evolutionary approach to the equilibrium, it will be useful to consider a special case. For simplicity, let us focus on the stage early in the process of evolving sexual dimorphism, when the variances the two sexes are equal,  $P_m = P_f = P$  and  $G_m = G_f = G$ , and the two sexes experience similar strengths of stabilizing selection,  $\omega_m = \omega_f = \omega$ . It will also be useful to define the sexual average as  $a = \frac{1}{2}(\bar{z} + \bar{y})$  and sexual dimorphism as  $d = \bar{z} - \bar{y}$ . These new traits,  $a$  and  $d$ , constitute a 90 deg rotation of the axes in Fig. 12.x. Substituting these definitions of  $a$  and  $d$  into (12.03) and (12.02), we obtain equations for the evolution of the sexual average and sexual dimorphism

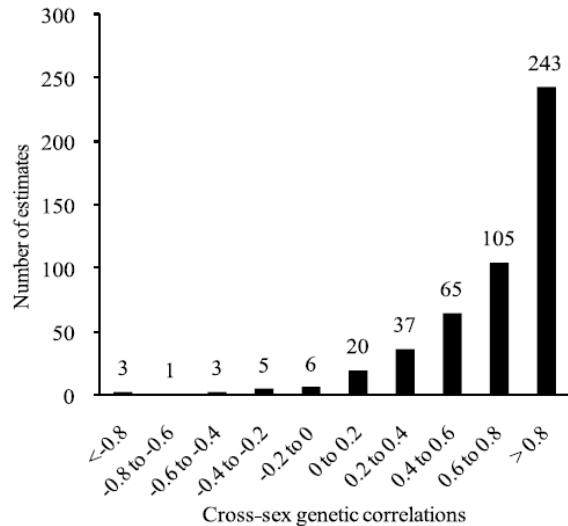
$$\begin{aligned} \Delta a &= \frac{1}{2}(G + B) \left\{ \frac{\frac{1}{2}(\theta_m + \theta_f) - a}{\omega + P} + \frac{1}{2}(c_m + c_f) \right\} \\ \Delta d &= \frac{1}{2}(G - B) \left\{ \frac{(\theta_m - \theta_f) - d}{\omega + P} + (c_m - c_f) \right\}. \end{aligned} \quad (12.05)$$

We see from (12.05) that the genetic covariance between the sexes,  $B$ , amplifies evolution of the sexual average but impedes evolution of sexual dimorphism ... {refer to rapid evolution of  $a$  followed by slow evolution of  $d$  ... departure from optima ... long period of maladaptation}

Empirical studies reveal a striking pattern of high genetic correlations between the sexes for homologous traits ... {Fig 12.4 = Figure from Poissant et al. 2009, n=488 estimates}

{need an empirical example of evolving sex dimorphism; check out Fairbairn & Prestesoi water strider papers ...}

{paragraph pointing out that above is a simplified, 2-trait portrayal of a model that allows multiple traits in each sex ... summarize some of the main conclusions from that more general model}



### 12.3 The balance between drift and stabilizing selection on a Gaussian landscape with a stationary peak.

We now consider the effect of finite population size in opposing the effects of stabilizing selection ...

... multivariate version of drift-selection balance with a Gaussian surface imposing the selection (section 11.1) ... Hansen & Martins (1997)

Each generation, the overall change in the lineage mean is the sum of two contributions: drift, which tends to move the trait mean away from the optimum, and stabilizing selection, which tends to pull the trait mean back towards the optimum,

$$\bar{z}_{t+1} = \bar{z}_t + G(\omega + P)^{-1}[\theta - \bar{z}_t] + N(0, G/N_e),$$

where the last term represents a random draw from a normal distribution with mean of zero and a variance-covariance matrix of  $G/N_e$ . We recognize the second term on the right as a response to directional selection,  $G\beta$ . At generation  $t$ , the lineage means are normally distributed about a mean vector of zeros with a variance-covariance matrix given by

$$Var(\bar{z}_t) = \frac{(\omega + P)}{2N_e} \left\{ 1 - \exp[-2tG(\omega + P)^{-1}] \right\}.$$

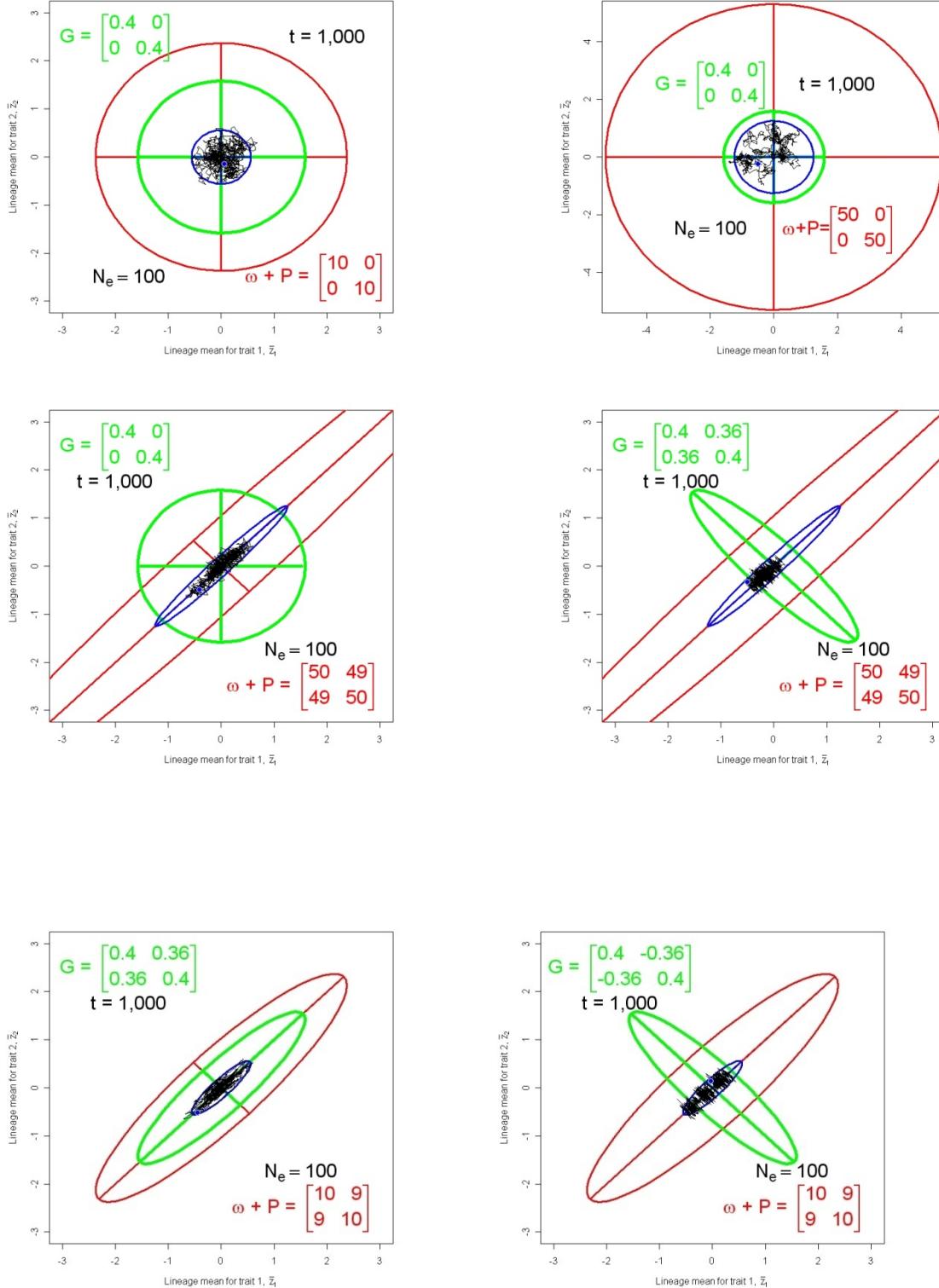
The first term involving  $\omega + P$  represents the balance between drift and stabilizing selection, while the second represents responses to directional selection. In the last term, {}, we see exponential decay in the influence of the  $G$ -matrix, so that the dispersion matrix achieves a limiting value given by

$$Var(\bar{z}_\infty) = \frac{(\omega + P)}{2N_e}$$

Consequently, on long time scales, the probability distribution of lineage means is normal and converges on,

$$\Phi(\bar{z}_\infty) = \sqrt{(2\pi)^{-n} |Var(\bar{z}_\infty)^{-1}|} \exp\{-(\bar{z} - \theta)Var(\bar{z}_\infty)^{-1}(\bar{z} - \theta)\}$$

Notice that while the AL appears in the last two equations, the G-matrix does not, indicating that the long term pattern of lineage means is shaped by the AL, not by the G-matrix.

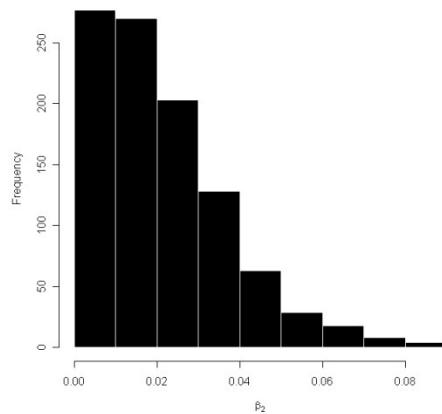


**Figure 12.5** {show only the variable parameters on the panels, not the constant ones; ie, no  $t$  and  $N_e$ } The role of the adaptive landscape in shaping the results of drift-selection balance. The 95% confidence ellipse for genetic values is shown in green and represents the G-matrix. The adaptive landscape is bivariate stabilizing in all cases and is represented by its 25% confidence ellipse in red. Such a low level ellipse must be employed because, the limits of the 95% ellipse would be far outside the scale limits of the figure. The limiting pattern for dispersion of lineage means,  $Var(\bar{z}_\infty)$ , is represented by its 95% confidence ellipse in blue. In each panel the black line shows the results of 1000 generations of simulated evolution of the lineage mean under the conditions specified in each panel. (e [should be a]) Weak stabilizing selection with no selectional correlation and no genetic correlation. (f [should be b]). Very weak stabilizing selection with no selectional correlation and no genetic correlation. Notice the change in scale. (a [should be c]). Very weak stabilizing selection with strong positive selectional correlation and no genetic correlation. (b [should be d]) Very weak stabilizing selection with strong positive selectional correlation and a strong negative genetic correlation. (c [should be e]) Weak stabilizing selection with strong positive selectional correlation and a strong positive genetic correlation. (d [should be f]) Weak stabilizing selection with strong positive selectional correlation and a strong negative genetic correlation.

The adaptive landscape exerts a strong influence on the outcome of drift-selection balance. Figure 12.2 illustrates the relevant case of bivariate stabilizing selection, showing that even very weak stabilizing selection strongly constrains the extent to which lineage means depart from the immediate vicinity of the adaptive peak. To appreciate the remarkable constraint exerted by the AL compare Fig. 12.2 with Fig. 8.1 which shows the 95% confidence ellipse for lineage means evolving by drift alone. Those ellipses are considerably larger than the comparable ellipse for the G-matrix. In contrast, even very weak stabilizing selection constrains evolving lineage means so that they occupy a bivariate area that is a fraction of the size of the ellipse that represents genetic values dispersed around the bivariate mean (Fig. 12.2a, b). The AL also strongly constrains the shape of the lineage dispersion. Those dispersion patterns (shown as blue ellipses), always mirror the shape of the AL and are completely independent of the shape of the G-matrix, as shown by (12.x). Finally, notice that a particularly favorable case for lineage dispersion is presented in Fig. 12.2 by using a very small effective population size ( $N_e = 100$ ). For more realistic sizes, say  $N_e$  in the range 500 to 5,000, the lineage dispersion ellipses would be minuscule and the paths of simulated lineage paths could not be visually resolved. It is no exaggeration to say that, in the absence of peak movement, even very weak stabilizing selection eliminates the possibility of appreciable radiation among lineages.

Is the process of drift-selection balance capable of producing the kind of distribution of directional selection gradients that we see in nature? ...

**Fig. 12.6** ... the process produces the right range of  $\beta$ -values and the distribution is almost the right shape ... a shortage of values very close to zero, corresponding to mean very close to the optimum ... despite this discrepancy, the overall distribution is much closer to observation than one produced by approach to the optimum from afar (Fig. 12.2) ...



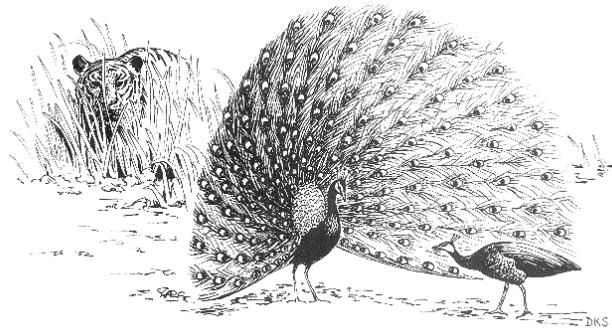
## 12.4 Hill climbing with two adaptive peaks

... another role for genetic correlation ... can affect which of two peaks is climbed ... {pull in illustration from Kirkpatrick or Schluter or create one = Fig. 12.7}

## 12.5 The joint evolution of sexual preference and sexually-selected traits

Darwin (1859, 1874) proposed a special process, sexual selection, to account for the evolution of traits that aid males in obtaining mates by sexual persuasion of females or by physical offense and defense in contests with sexual rivals. In Darwin's view, ordinary viability selection could not account for the elaboration of such traits, and indeed that selective agency would often act in the opposite direction, toward diminution. While sexual preference of females for elaborate male traits seemed inevitable to Darwin, it remained controversial for decades into the 20<sup>th</sup> century. Fisher (1915, 1930), however, sketched a model for Darwin's process of mate choice that became famous for its prediction of runaway evolution of female preference and male trait. In this section, we summarize Lande's (1981) model of the Darwin-Fisher proposal, with some simplifications in the interest of getting quickly to the main conclusions.

Lande (1981) modeled the joint evolution of two traits with sex-limited expression: a male trait, that we will call the ornament (e.g., tail size in peacocks), and a female trait, sexual preference for the male ornament (Fig. 12.8). This model has the interesting feature that evolution results in a line of equilibria, rather than a single unique outcome. The line of equilibria is a consequence of specifying two types of selection on the male ornament: viability selection represented by the tiger in Fig. 12.9, and mating preferences, represented by the peahen in the same figure.



Let  $z$  be the phenotypic value of a normally distributed male ornament. Before selection the mean of  $z$  is  $\bar{z}$  and its variance is  $P_m$ . The ornament is affected by many autosomal genes of small effect so that its breeding value is normally distributed with genetic variance  $G$ . Each generation Gaussian viability selection acts on the ornament in a stabilizing mode with an intermediate optimum  $\theta_m$ . This viability selection changes the distribution of  $z$  so that after selection its mean and variance are

$$\bar{z}^* = (\bar{z}\omega + \theta P_m) / (\omega + P_m) \quad (12.06)$$

$$P_m^* = \omega P_m / (\omega + P_m), \quad (12.07)$$

where  $\omega$  is the ‘variance’ of the viability function, so that the selection differential due to viability selection is

$$\bar{z}^* - \bar{z} = (\theta - \bar{z}) / (\omega + P_m)$$

Each generation, following viability selection, the population of surviving males is subjected to sexual selection by females that choose their mates on the basis of the male ornament. Female preference,  $y$ , is a normally distributed trait with phenotypic mean  $\bar{y}$  and variance  $\tau^2$ . We suppose that female preference is affected by many genes, so that breeding values are normally-distributed with variance  $H$ . Female choice

of males occurs in the following way. Each female is characterized by a Gaussian mate preference function with an optimum at  $y$  and a ‘variance’ of  $v^2$ ,

$$\psi(z|y) \propto \exp\{-(z-y)^2/2v^2\}.$$

In other words, a female is most likely to mate with an encountered male if his ornament matches her preference, so that  $z = y$ , and her tendency to mate falls off as a Gaussian function as his ornament deviates from that optimum,  $y$ . Averaging these preference functions over the population of females, we obtain a Gaussian function that gives us the overall probability of mating as a function of male ornament value,  $z$ ,

$$\psi(z) = b \exp\{-(z - \bar{y})^2 / 2(\tau^2 + v^2)\}$$

where  $b = \dots$ . After sexual selection by a subset of female with preference  $y$ , the ornament mean  $\bar{z}^*$  will be shifted by an amount

$$(y - \bar{z}^*)P_m*/(v^2 + P_m^*) .$$

Averaging these shifts over the entire female population yields the selection differential due to sexual preference,  $\bar{z}^{**} - \bar{z}^*$ . The total selection differential is the sum of viability and sexual differentials, giving a total selection gradient for the ornament

$$P_m^{-1}s = P_m^{-1}[(\bar{z}^{**} - \bar{z}^*) + (\bar{z}^* - \bar{z})] = P_m^{-1}(\bar{z}^{**} - \bar{z})$$

$$P_m^{-1}s \approx \frac{\bar{y}/\alpha - (1+1/\alpha)\bar{z} + \theta}{\omega},$$

where  $\alpha \equiv v^2/\omega$ .

Turning to females, let us assume that a female’s fitness (progeny count) is unaffected by her mate choice, with the consequence that no selection acts on female preference,  $y$ . As a result, average preference,  $\bar{y}$ , will evolve only as a correlated response to selection on the ornament. We realize from earlier results (xxxx), that such a correlated response will depend on a genetic covariance between the two traits,  $B$ , a genetic covariance between the sexes. This genetic covariance is unlikely to be a consequence of pleiotropy, because in general it is difficult to imagine a contributing locus that affects a male ornament will also have a pleiotropic effect on female preference, or vice versa. On the other hand,  $B$  could reflect linkage disequilibrium. Mate choice in our model leads to assortative mating between the  $z$  and  $y$  phenotypes as well as selection that favors particular combinations of  $z$  and  $y$ . Lande (1981) reports a submodel showing that linkage disequilibrium does result from this combination of assortative mating and sexual selection and accumulates across loci to constitute a genetic covariance between the sexes,  $B$ . Furthermore, the primary determinate of the magnitude of  $B$  is ratio of mutation rates affecting female preferences versus the male ornament {a little more elaboration here, see Lande section}. Putting these results together with an expectation that  $G$  and  $H$  will equilibrate in mutation-selection balance, we can conclude that the G-matrix for our two characters,

$$\begin{bmatrix} G & B \\ B & H \end{bmatrix},$$

is likely to be stable if stabilizing selection on the preference is weak ( $\omega \gg P_m$ ).

Using our standard expression for the response to two traits to selection (e.g., 12.00) and recalling that we have not allowed direct selection on  $y$ , we have

$$\Delta\bar{z} = \frac{1}{2}GP_m^{-1}s \quad (12.05a)$$

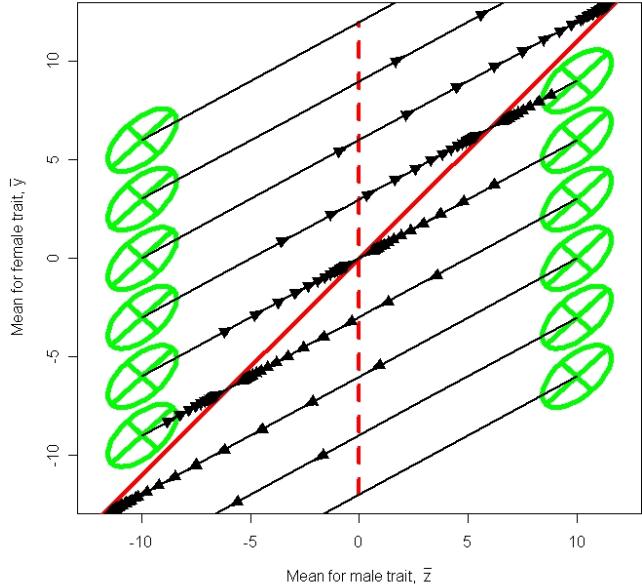
$$\Delta\bar{y} = \frac{1}{2}BP_m^{-1}s \quad (12.05b),$$

...  
To find the equilibrium, we set the total selection gradient equal to zero and find that the equilibrium is a line

$$\bar{y} = (\alpha + 1)\bar{z} - \alpha\theta. \quad (12.13)$$

In other words, at any position along the equilibrium line, the force of viability selection pulling  $\bar{z}$  towards the optimum  $\theta$  is exactly balanced by a force of sexual selection in the opposite direction (Fig. 12.5).

**Figure 12.10.** Lande's 1981 model for the joint evolution of a male ornament and female mating preferences based on that ornament. The stable case is illustrated. The solid red line is the line of equilibria. The dashed red line shows the position of the viability selection optimum,  $\theta$ , for the male trait. Solid black lines show evolutionary trajectories for 12 starting positions, with arrows spaced every generation. The green ellipses at the starting points are G-matrices.  $G=H=0.4$ ;  $B=0.24$ ;  $\omega=4$ ,  $\alpha=0.1$ ,  $r^2=P_m=1$ . The similarity between  $g_{max}$  and the line of equilibria in this figure is a coincidence.



To analyze the stability properties of this equilibrium, it is useful to shift the coordinate system by a constant in the  $\bar{z}$  dimension {illustrate this shifted coord system on a fig} by defining

$$\tilde{z} = \bar{z} - \theta/(1+1/\alpha) \quad (12.14a)$$

$$\tilde{y} = \bar{y} \quad (12.14b)$$

In this new coordinate system, the selection gradient is

$$\beta = P_m^{-1}s = \frac{\tilde{y}/\alpha - \tilde{z}(1+1/\alpha)}{\omega} \quad (12.15)$$

and consequently the line of equilibria is

$$\tilde{y} = (\alpha + 1)\tilde{z} \quad (12.16)$$

Substituting (12.15) into (12.05), we obtain a simple expression for evolution in our new coordinate frame,

$$\begin{bmatrix} \Delta\tilde{z} \\ \Delta\tilde{y} \end{bmatrix} = \frac{1}{2\alpha\omega} \begin{bmatrix} -(\alpha+1)G & G \\ -(\alpha+1)B & B \end{bmatrix} \begin{bmatrix} \tilde{z} \\ \tilde{y} \end{bmatrix} = \frac{1}{2\alpha\omega} \begin{bmatrix} -(\alpha+1)G\tilde{z} + G\tilde{y} \\ -(\alpha+1)B\tilde{z} + B\tilde{y} \end{bmatrix}. \quad (12.17)$$

From (12.05) we see that  $\Delta\bar{y}/\Delta\bar{z} = B/G$ , indicating that evolution will occur along straight line trajectories with a slope given by the genetic regression  $B/G$ . From (12.17), we obtain the same result. Taking the eigenvalues of the matrix on the left side of (12.17), we find that one eigenvector corresponds to  $B/G$ , and its eigenvalue is

$$\lambda = [B - (\alpha+1)G]/(2\alpha\omega). \quad (12.18)$$

Consequently, we can conclude that after  $t$  elapsed generations, the vector of trait means evolving along a  $B/G$  trajectory will have changed by an amount  $(1+\lambda)^t$ . If  $B/G < (\alpha+1)$ , then  $\lambda$  is negative, and evolving populations approach the line of equilibria with ever decreasing speed; the equilibrium is stable (Fig. 12.5). But if  $B/G > (\alpha+1)$ ,  $\lambda$  is positive, and populations evolve away from the line of equilibrium at ever increasing speed (as proposed by Fisher 1958); the equilibrium is unstable. The other eigenvector is the line of equilibrium itself with an associated eigenvalue equal to zero, indicating responses to selection cease on that line. The simplest change in ecology that could trigger the unstable, runaway case is a relaxation in stabilizing selection (an increase in  $\omega$ ) so that  $B/G > (\nu^2/\omega) + 1$ .

So far, we have considered only deterministic evolution of preference and ornament, ignoring the possibility of drift. The existence of a line of equilibria invites us to consider drift because, especially in the stable case, the bivariate mean might drift up and down that line. We need to consider situations in which effective population size is relatively large ( $N_e \geq 500$  or 1000), so that the G-matrix is likely to be relatively constant. Under these conditions results from section (xx) tell us that random sampling of parents will increases the variance and covariance of replicate populations means by a constant amount each generation,

$$Cov(\bar{z}, \bar{y}) = \frac{1}{N_e} \begin{bmatrix} G & B \\ B & H \end{bmatrix}. \quad (12.19)$$

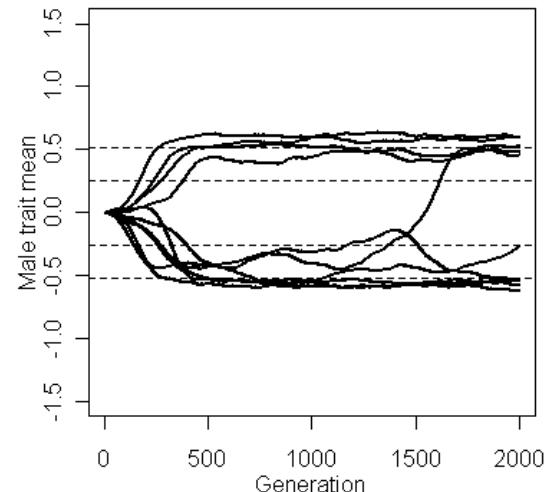
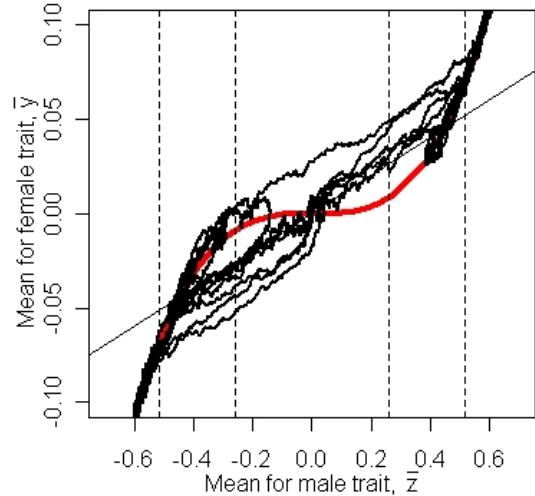
After  $t$  generations have elapsed, Lande (1981) shows that in the stable case the dispersion among replicate means will be

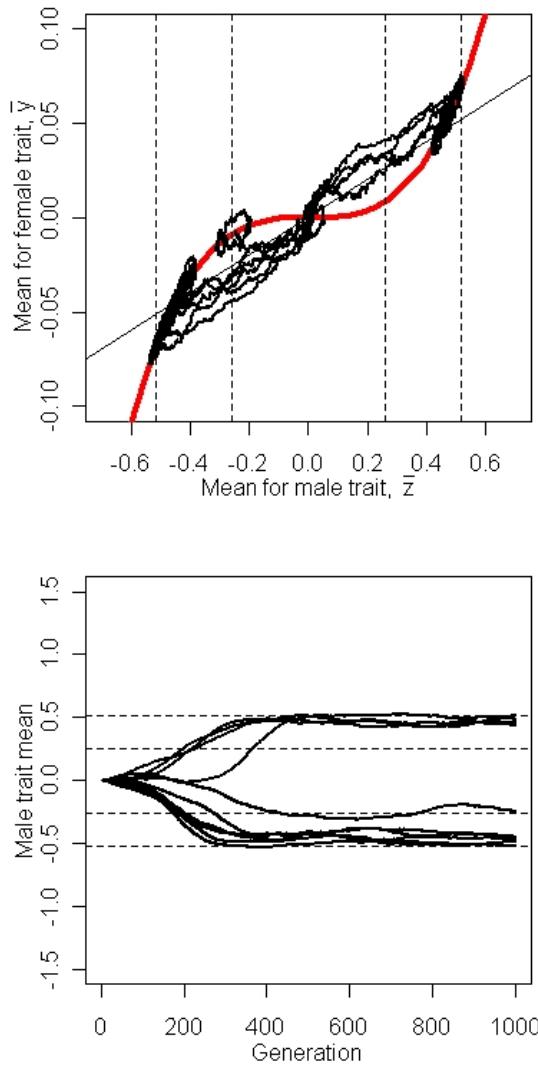
$$D(t) \approx K + \frac{H(1-r_g^2)t}{N_e(\alpha+1-B/G)^2} \begin{bmatrix} 1 & \alpha+1 \\ \alpha+1 & (\alpha+1)^2 \end{bmatrix} = \begin{bmatrix} Var(\bar{z})_t & Cov(\bar{z}, \bar{y})_t \\ Cov(\bar{z}, \bar{y})_t & Var(\bar{y})_t \end{bmatrix}, \quad (12.20)$$

where  $K$  is a constant matrix and  $r_g = B/\sqrt{GH}$ . Thus, genetic variance for preference,  $H$ , which played no role in deterministic responses to selection, plays a major role in the drift of both preference and ornament. This result makes intuitive sense. The preference distribution is not under direct selection and is free to drift along the line of equilibria. As it drifts, that preference distribution exerts directional selection on the male ornament distribution, so that the drift of  $\bar{z}$  is paced by the inheritance of  $y$ , i.e., by  $H$ .

Simulations by Uyeda et al. (2009) confirm Lande's (1981) interpretation that drift along the line of equilibria can produce rapid evolutionary diversification, even with  $N_e$  as large as 5,000-10,000. Those same simulations reveal that  $\bar{z}$  and  $\bar{y}$  are extremely close together as a lineage drifts along the equilibrium line, so that departures from the line are nonobvious, and that the two means often switch positions, so that sometimes  $\bar{y} > \bar{z}$  and sometimes  $\bar{z} > \bar{y}$ .

Iwasa & Pomiankowski (1995) have described a process that yields perpetual evolution of a male ornament and female mating preference ... differs from Lande's (1981) model in a couple of important respects ... viability selection on the male trait is a fourth-power function ...  $\exp\{c\bar{z}^4\}$ , where  $c$  is a constant, so that the viability function is more flat-topped than a Gaussian or quadratic function. The optimum trait value with respect to viability selection is arbitrarily set to 0. ... female mating preference for a particular male is an exponential function of his trait value ...  $\exp\{a(z - \bar{z})\bar{y}\}$ , where  $a$  is a constant ... selection acts on female mating preferences, so that female fitness with respect to her preference value  $y$  is  $\exp\{-by^2\}$ . Ordinarily, the effect of selection on female preferences is to collapse a line of equilibria, as in Fig. 12.x, to a point, so that in the stable case, the joint evolution of ornament and preference has a single unique outcome (Pomiankowski et al. 1991a, b). In the present model, however, when the assumption of selection on preferences is combined with a flat-topped, viability selection function for the male trait, the combined effect is to produce a curve of equilibrium that is a cubic function. This cubic function has the remarkable property of producing a stable limit cycle in a population of infinite size (Iwasa & Pomiankowski 1995). This regular cyclical behavior of the model disappears in populations of finite size even if they are very large (Fig. 12.11), but nevertheless evolution appears to consist of perpetual swings between alternative stable zones on either side of the viability optimum,  $\bar{z} = 0$ .





**Figure 12.11.** Perpetual evolution of a male ornament and female mating preferences. In this model the line of equilibria, the red curves in the upper panels, is a cubic function. This line of equilibria consists of stable and unstable zones. The unstable zone around the inflection of the red curve is denoted by the inner pair of dashed lines. The transition to flanking stable zones is located between the pairs of dashed lines. The evolutionary trajectories of two sets of 10 replicate lineages are shown with origination at  $\bar{z} = 0, \bar{y} = 0$ . Populations off the line of equilibria evolve deterministically along trajectories with a slope of  $B/G$ ; one such line (black) is shown through the origin. (a) On a 1,000 generation time scale, the tendency of lineages to stall within the zones of stability is apparent. (b) On a 2,000 generation time scale, diversification within the stable zones has continued, and one lineage has evolved across the unstable zone, from one stable zone to the other. Parameter values are:  $N_e=10^6$ ,  $G=H=0.5$ ,  $B=0.05$ ,  $a=0.4$ ,  $b=0.001$ ,  $c=0.05$ ,  $u=0$ , where  $u$  represents biased mutation.

In Fig. 12.11b, a single lineage makes the transit between alternative stable zones, but on a longer time scale, such transits are relatively common. Alternatively, with  $N_e$  smaller but still large (e.g., 10,000), transits are common on a time scale  $\geq 1,000$  generations.

In several respects the Iwasa & Pomiankowski 1995 model should be considered a work in progress. For one thing, because of the model's parameterization, we do not know whether its assumptions about forms and intensity of selection on the two traits match empirical reality. Secondly, we do not know whether its signature feature, limit cycles, depends on a narrow choice of parameter values or persists over a broad range. Finally, the function that specifies the distribution of trait means as a function of time and  $N_e$  is unknown. Nevertheless, the model emphasizes the essential point, discussed below, that a variety of evolutionary outcomes are possible when two traits have interactive effects on fitness.

The 30+ models that have been based on the framework established by Lande (1981) have explored a variety of alternative assumptions and established some important generalizations (Mead & Arnold 2004) ... a variety of outcomes are possible, depending mainly on assumptions about selection (Fig. 12.6) ... alternative assumptions have been used to capture the essential features of 'good genes', sexual conflict, costly female choice, ...despite the variety of possible outcomes, all models reveal both stable and unstable outcomes ... most models have not explored drift ... refer to explorations described by Uyeda et al 2009 ... end with a statement about implications for speciation.

{In a final paragraph for this section, connect up these modeling efforts with the empirical literature on sexual selection ... estimates of the heritability of male traits and female preference (find the most recent review) ... estimates of the genetic correlation between preference and trait (find the most recent review) ... tests for various kinds of selection ... finally, comment on the unsatisfactory state of tests for alternative models}

### 12.9 The evolution of sexual isolation

Lines of equilibria and stable limit cycles imply that diversification by sexual selection could lead to speciation, perhaps easily and rapidly. To verify those implications and predict rates of speciation, we need a model that connects sexual selection to sexual isolation, and speciation. One way to proceed is to make those connections using Lande's (1981) model, outlined in the last section (Arnold et al. 1996). In particular, we can use (12.0x) to specify the probability that a randomly drawn female from one population,  $A$ , will mate with a randomly drawn male from another population,  $B$ ,

$$\pi_{AB} = c \exp\{-(\bar{z}_B - \bar{y}_A)^2 / 2\sigma^2\}, \quad (12.21)$$

where  $c$  is a constant less than 1 but greater than 0, and the variation parameter  $\sigma^2 = \nu^2 + \tau^2 + P_m$  is assumed to be the same in both populations. This same expression can be used to specify the probabilities of mating,  $\pi_{AA}$  and  $\pi_{BB}$ , when mating partners are drawn from the same population or from different populations in opposite order, female from  $B$  and male from  $A$ ,  $\pi_{BA}$ . These same four probabilities can be estimated in an experimental assay of sexual isolation. In such an assay, mating partners are drawn at random from two different populations, four kinds of sexual encounter are staged and replicated, mating success is scored for each kind, and the four probabilities of mating are estimated directly from those data. One common and useful measure of sexual isolation that summarizes such data is called the *index of joint isolation*,

$$JI = \pi_{AA} + \pi_{BB} - \pi_{AB} - \pi_{BA}.$$

This index ranges from +2 when ... to -2 when ...

To model the evolution of  $Jl$ , we need to consider a representative pair of populations,  $A$  and  $B$ , that have evolved to the line of stable equilibrium in Lande's (1981) model and are now drifting along it. To proceed, it will help to make some simplifying assumptions. First, we note (and confirm by simulation) that the difference between preference and ornament means within a population will be negligible relative to the expected between-population difference in means. Under this assumption, from (12.xx),  $\pi_{AA} = \pi_{BB} = 1$  and  $\pi_{AB} = \pi_{BA}$ . The distribution of  $\pi_{AB} = \pi_{BA}$  among replicate pairs of populations is then a function of the corresponding distribution of pairs of ornament means (Fig. 12.7), which we can determine using (12.xx). Pursuing this course, we find that the function giving the expected value of  $Jl$  as a function of time is not simple, but it can be evaluated numerically or the process can be simulated (Uyeda et al. 2009). Some sample paths are shown in

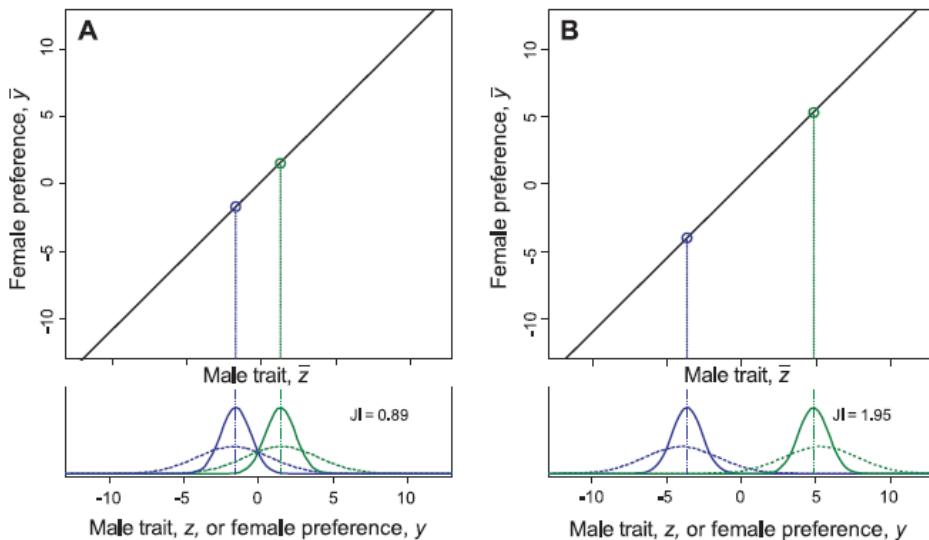
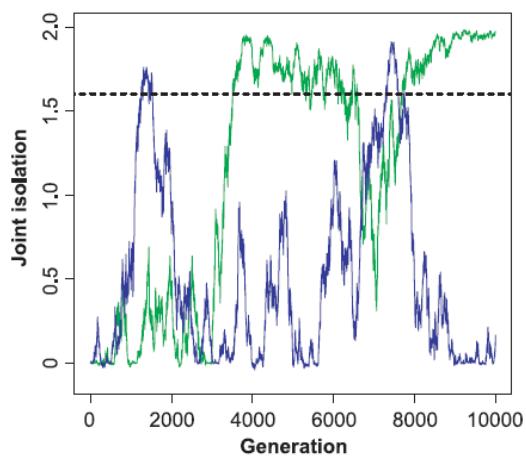
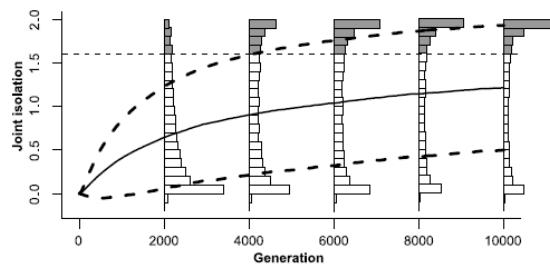
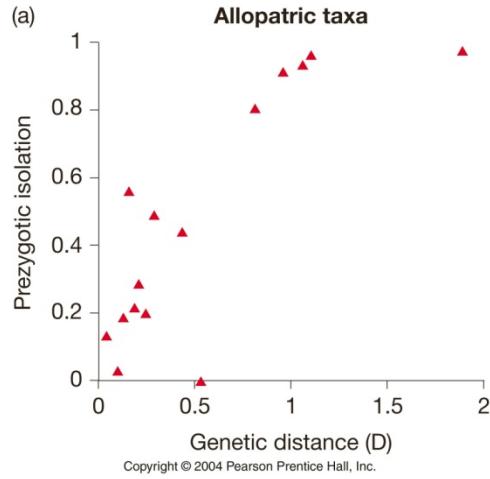


Fig 12.7. Simulations reveal that the evolutionary trajectory of  $Jl$  for any pair of populations is often chaotic, as we might expect from the overwhelming influence of drift. Instead of progressing steadily towards increasing isolation, individual trajectories often alternate between episodes in which isolation waxes and wanes (Fig. 12.8). On the other hand, when we follow the histories of many replicate pairs of populations, we do see a trend towards increasing isolation (Fig. 12.9). Nevertheless, the distribution of  $Jl$  is characteristically bimodal during intermediate stages, until finally the mode becomes strong sexual isolation ( $Jl \approx 2$ ), typically after some thousands of generations.



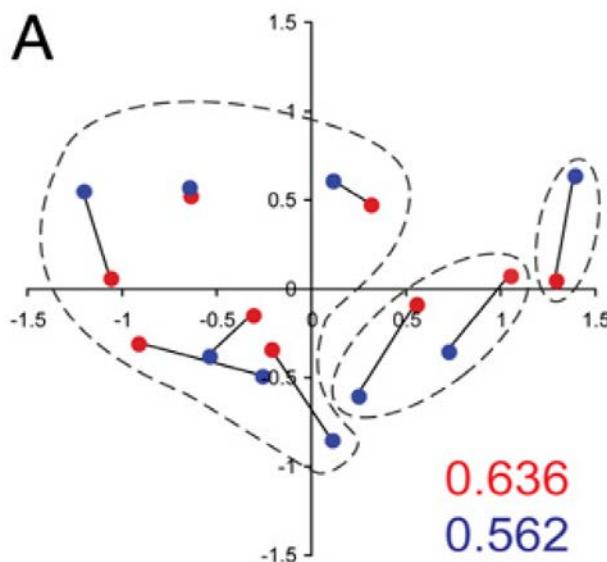
{go on to discuss influence of  $N_e$ , etc. ...; relate this model to the evolution of  $Jl$  in drosophila and desmogs, perhaps producing histograms of  $Jl$  that can be compared to the fig on the left; bring in a figure from Coyne & Orr 1997 showing  $Jl$  as a function of

genetic distance and comment especially on the great spread in  $J_1$  as short genetic distances, changing asymptotically to high values at long Nei distance}



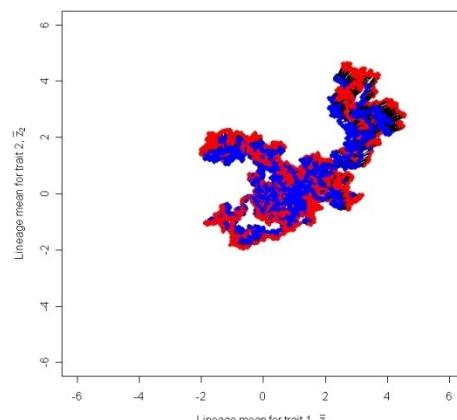
### 12.10 The dimensionality of mate choice and sexual isolation

... from a variety of perspectives it seems likely that mate choice is based on multiple traits ... the possibility that sexual selection involves multiple traits sits in stark contrast to the modeling preoccupation with single traits ... Pomiankowski & Iwasa ... Iwasa & Pomiankowski (1993, 1994) are the exception ... they found that ... Hohenlohe & Arnold (2010) took another approach by ... evidence that sexual isolation based on two or more trait dimensions in drosophila, cichlid fishes and plethodontid salamanders ...



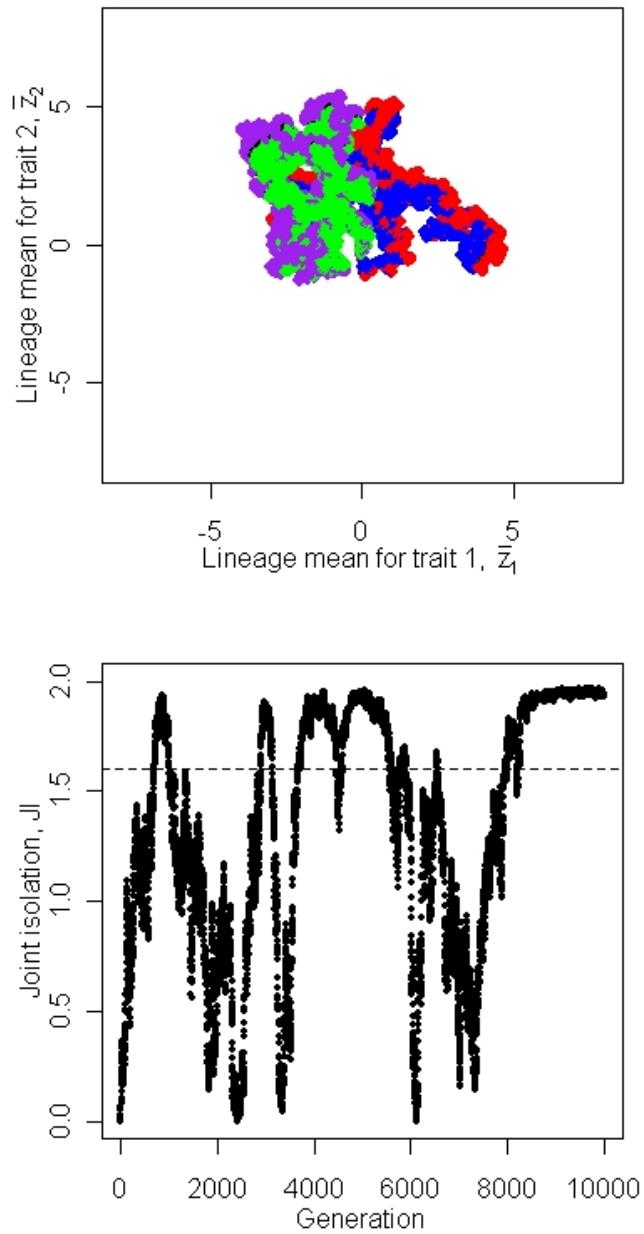
**Fig. 12.11.** Latent axes for male and female traits in *Desmognathus* salamanders, inferred from a survey of sexual isolation. Male mean (blue) and female means (red) from the same population are connected by solid lines. Hohenlohe & Arnold 2010.

**Figure 12.12.** Joint evolution of two male traits and female mating preferences based on them on a stable plane of equilibria. Male means are shown in blue and female means are shown in red. Mean in the same generation are connected by solid black lines, but are often not visible because the



means of the two sexes are close together. 10,000 generation of interactions between drift and selection.  $N_e=500$ . The genetic variance of male and female traits is 0.4. Traits are genetically uncorrelated within sexes. Genetic correlations between the sexes are 0.6 for traits with the same subscript and 0 otherwise. For both male traits,  $\omega=4$  (with no correlational selection and  $P_m=1$ ; for both both female traits,  $v^2=0.4$  (with no correlational preferences),  $\tau^2=1$ , and  $\alpha=0.1$ . The optimum or the male traits is situated at 0, 0.

**Figure 12.13** Sexual isolation evolving as a function of the joint evolution of two male traits and female mating preferences based on them in two, independent lineages. (a) Joint evolution of bivariate male and female traits in two lineages. Same conventions and parameter values as in Fig. 12.12 except that one pair of means is colored blue and red, and the other pair is colored purple and green. The second lineage partially obscures the history of the first lineage. (b) Time course for joint isolation,  $J_I$ . The horizontal dashed line shows the lower level of  $J_I$  that is characteristic of sympatric, isolated species.



## Chapter 13: Trait Evolution on Dynamic Adaptive Landscapes: Stochastic Models

© Stevan J. Arnold 2016

**Overview.-** Although a variety of models have been proposed to explain adaptive radiations in a single trait, a common feature of relatively successful models is that the peak of a hill-shaped adaptive landscape moves, but the shape of the hill does not change. Although many of these models fail to explain temporal patterns in large data set assembled by Philip Gingerich and others, models in which the optimum fluctuates with Brownian motion can approximate data from adaptive radiations. Models that come closest to accounting for the data are ones in which the optimum moves rarely, but is capable of substantial excursions (0-6 phenotypic standard deviations). These results direct our attention to the problem of formulating and testing alternative models for the peak controller process (PCP).

{introductory paragraph here, sketching the rationale for the approach (see section in Estes & Arnold 2007), anticipating the main results that follow, and using Fig. 13.0 = cartoon of the models of peak movement as a guide ...}

### 13.0 Brownian motion of the adaptive peak

In Chapter 11 we showed that the drift of populations of finite size on an adaptive landscape with a single stationary adaptive peak does not explain the diversification commonly seen in evolving clades. The basic problem is that a stationary peak does not permit enough variation among lineage means. We can imagine solving this problem by allowing the peak to move in some pattern. In this section we explore Brownian motion of the peak and evaluate the diversification produced by that motion.

By Brownian motion of the optimum we mean that the current position of the optimum deviates from its position in the preceding time interval and by an amount that is a normally distributed random variable:  $\theta_{t+1} = \theta_t + \varepsilon_\theta$ , where the last term is random variable drawn from a normal distribution with a mean of zero and variance of  $\sigma_\theta^2$ . We want to determine how a set of replicate populations will evolve in response to this mode of peak movement. If we focus on a single evolving lineage, the change in the trait mean in a particular generation ( $t+1$ ) has two components

$$\Delta\bar{z}(t+1) = \left( \frac{\theta_t - \bar{z}_t}{\omega + P} \right) G + N(0, G/N_e),$$

a deterministic response (the first term on the right) that is proportional to the distance to the optimum and a stochastic, drift response that is exacerbated by small population size. The deviation of the optimum from its position in the previous generation is a draw from a normal distribution with a mean of zero and a variance of  $\sigma_\theta^2$ ,

$$\Delta\theta(t+1) = \theta_t + N(0, \sigma_\theta^2).$$

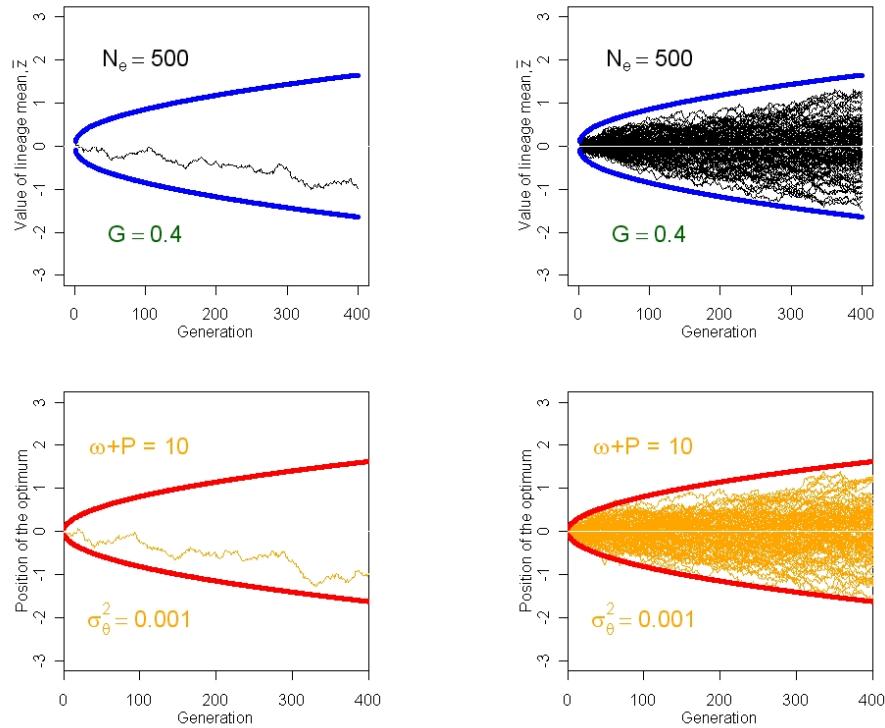
The distribution of the means of replicate lineages is normally distributed about the long term average position of the optimum with an ever-increasing variance (Fig. 13.1a),

$$Var[\bar{z}(t)] = \frac{\sigma_\theta^2 + \frac{G}{N_e}}{2a} \{1 - exp[-2at]\} + \sigma_\theta^2 t \left\{ 1 - 2 \frac{(1 - exp[-at])}{at} \right\},$$

where  $a = \frac{G}{\omega+P}$  (Hansen et al. 2008). In contrast, Hansen et al. (2008) find that the variance among replicates in the position of the optimum at generation  $t$  is

$$\text{Var}[\theta(t)] = \sigma_\theta^2 t.$$

We see that the expression for the variance for the variance among replicate means rapidly converges on  $\sigma_\theta^2 t$ , which is the variance in the position of the optimum. An example of lineage responses to Brownian motion of adaptive peaks is provided in Fig. 13.3. Using reasonable values for selection and inheritance, we see – as expected – that the dispersion of lineage means is largely a reflection of variation in the position of optima.



**Figure 13.3** Response of the lineage mean to an intermediate optimum that undergoes a random walk. Lower panels show the path of the optimum,  $\theta$ , through time. The 99% confidence limits of the optimum are shown with red lines. The upper panels show the paths of lineage means,  $\bar{z}$ . 99% confidence limits are shown with blue lines. (a) A single lineage mean closely tracks the moving optimum of its AL. (b) 100 replicate lineages track the moving optima of their ALs.

An example of lineage responses to Brownian motion of adaptive peaks is provided in Fig. 13.3. Using reasonable values for selection and inheritance, we see that the dispersion of lineage means is largely a reflection of variation in the position of optima.

### 13.1 White noise motion of the adaptive peak

We now turn to a more dramatic mode of peak movement, white noise motion. In this mode our model of the position of the optimum is  $\theta_{t+1} = \varepsilon_\theta$ , where, as before, the last term is random variable drawn from a normal distribution with a mean of zero and variance of  $\sigma_\theta^2$ . Notice that the optimum carries none of its

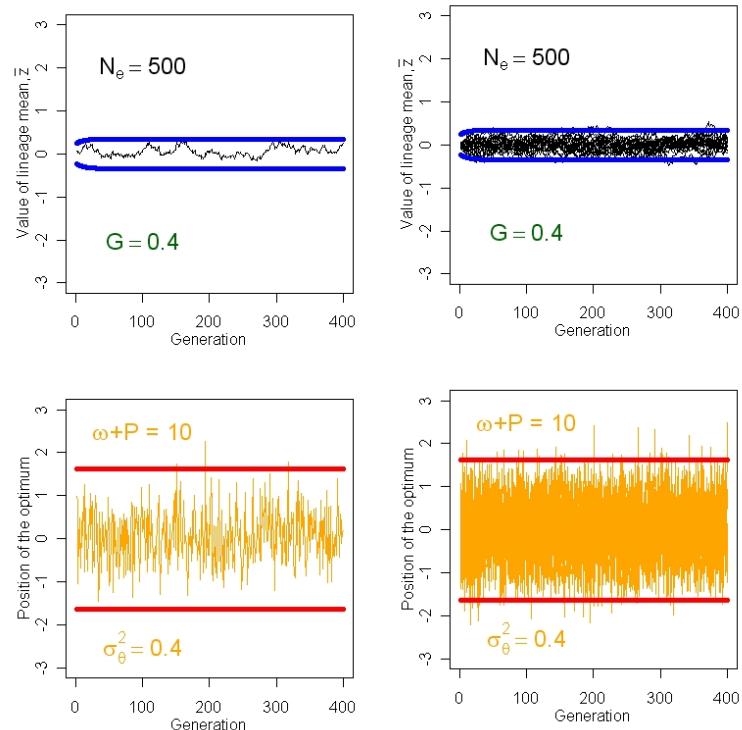
past history, as it does in the case of Brownian motion. Instead, the optimum undergoes more erratic motion. As in the last model, each generation of the trait mean of a lineage is affected by two factors, recent movement of the optimum and genetic drift of the trait mean. The trait means of replicate lineages are normally distributed with a expected value that corresponds to the expected value for the optimum, which we have assumed is zero. The variance among the trait means of replicate lineages at any particular generation  $t$  is

$$Var(\bar{z}_t) = \left[ \frac{G\sigma_\theta^2}{2(\omega+P)} + \frac{\omega+P}{2N_e} \right] \left\{ 1 - \exp \left[ -2 \left( \frac{G}{\omega+P} \right) t \right] \right\}$$

... the new term is the first one on the right, which represents the contribution from the trait mean as it tracks peak movement ... facilitated by genetic variance ... inhibited if stabilizing selection is weak ...

$$Var(\bar{z}_\infty) = \frac{G\sigma_\theta^2}{2(\omega+P)} + \frac{\omega+P}{2N_e}$$

... notice that the variance takes a constant value at equilibrium, in contrast to the Brownian motion case ... (Fig. 13.1b). A particular example of white noise motion and the evolution it evokes is shown in Fig. 13.5, which employs intermediate values for genetic variance and population size and imposes moderately weak stabilizing. Notice in Fig. 13.5a that the evolving mean fails to track the rapid, erratic motion of the optimum. As a consequence, the ensemble properties of many replicate populations are a much damped version of their moving optima (Fig. 13.5b). Another revealing contrast with Brownian motion is the general result that variance among replicate means rapidly achieves its asymptotic value instead of expanding endlessly. The consequence of this stationary variance is bounded evolution that reflects the principal features observed in large data sets (Fig.xxxxx) ... the problem of too little diversification ... how does the imposed variance in theta compare with observations in nature?



**Figure 13.5.** Simulations of lineages evolving in response to white noise motion of the optimum of the AL. The long term average position of the optimum is zero. The blue lines in the upper panels and red lines in the lower panels show the upper and lower 99% confidence limits (a) The trait mean of a single lineage (above) evolving in response the movement of the optimum of its AL. (b) The trait means of 20 replicate lineages (a) evolving in response to the movement of their ALs

### 13.2 Steady movement of the adaptive peak

In this model, in contrast to the models discussed so far, the trait mean shows a long term trend ... Our model of the position of the optimum is  $\theta_t = kt + \varepsilon_\theta$ , where  $k$  is a constant denoting the rate of deterministic change in the optimum. (Fig. 13.2a = showing both  $\theta$  and  $\bar{z}$  as a function of time) ... Notice that white noise motion of a stationary peak is a special case of this model ...

$$E[\bar{z}_t] = kt - k\left(\frac{\omega + P}{G}\right)\left\{1 - \exp\left[-\left(\frac{G}{\omega + P}\right)t\right]\right\}$$

... first term on the right represents response to directional selection elicited by peak movement ... the second term represents the negative contribution of stabilizing selection balance ... last term ... approach to equilibrium ... (Fig. 13.2b = based on Fig. 4 of Estes & Arnold 2007)

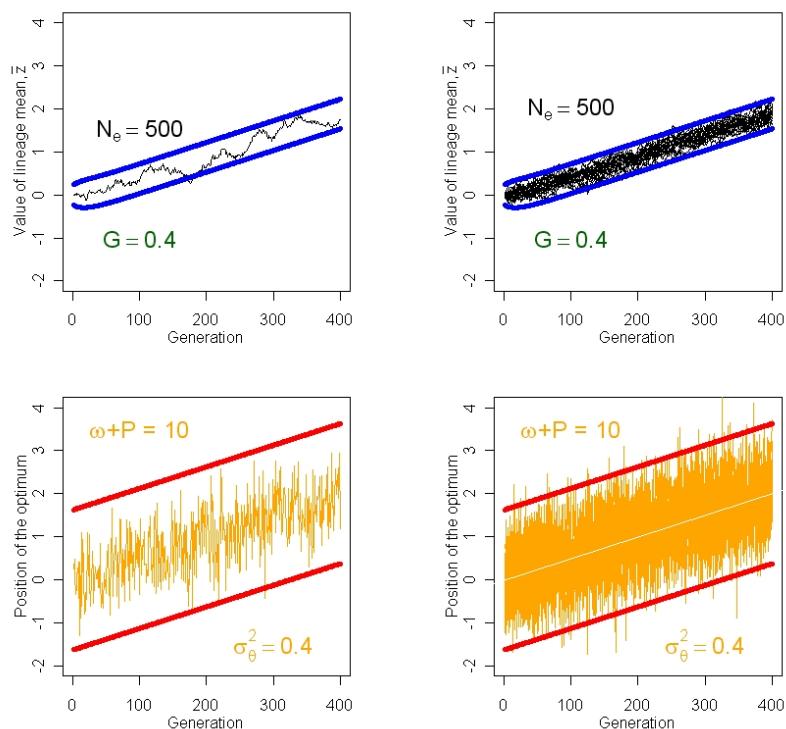
$$E[\bar{z}_\infty] = kt - k\left(\frac{\omega + P}{G}\right)$$

$$Var(\bar{z}_t) = \left[ \frac{G\sigma_\theta^2}{2(\omega + P)} + \frac{\omega + P}{2N_e} \right] \left\{ 1 - \exp\left[-2\left(\frac{G}{\omega + P}\right)t\right] \right\}$$

$$Var(\bar{z}_\infty) = \frac{G\sigma_\theta^2}{2(\omega + P)} + \frac{\omega + P}{2N_e}$$

An example of evolving populations evolving in response to steady movement of the adaptive peak is provided in Fig. 13.6. Notice that, as before, substantial white noise movement of the peak evokes only moderate excursions in the evolving trait mean.

**Figure 13.6** Simulations of lineages evolving in response to a steadily moving intermediate optimum with superimposed white noise fluctuations in position. Conventions as in Fig. 13.5. The optimum increases in value at a rate of 0.005 within-population standard deviation per generation ( $k =$



0.005). (a) The trait mean of a single lineage evolving in response to the moving optimum. (b) The trait means of 20 replicate lineages evolving in response to the same selection regime.

The Achilles heel of the steadily moving optimum model is that in the fullness of time the trait mean is bound to evolve outside the observed boundaries for divergence. For example, even if the optimum moves at a rate that would be undetectable in microevolutionary analyses ( $k=0.0002$ ), the average lineage will have diverged about 6 standard deviations after 30,000 generations (Fig. 13.7). Although this excursion past the boundary may not seem egregious, steadily moving optima lead to impossible levels of divergence on a time scales that exceed 100,000 generations, under all combinations of realistic parameter values (Estes & Arnold 2007).

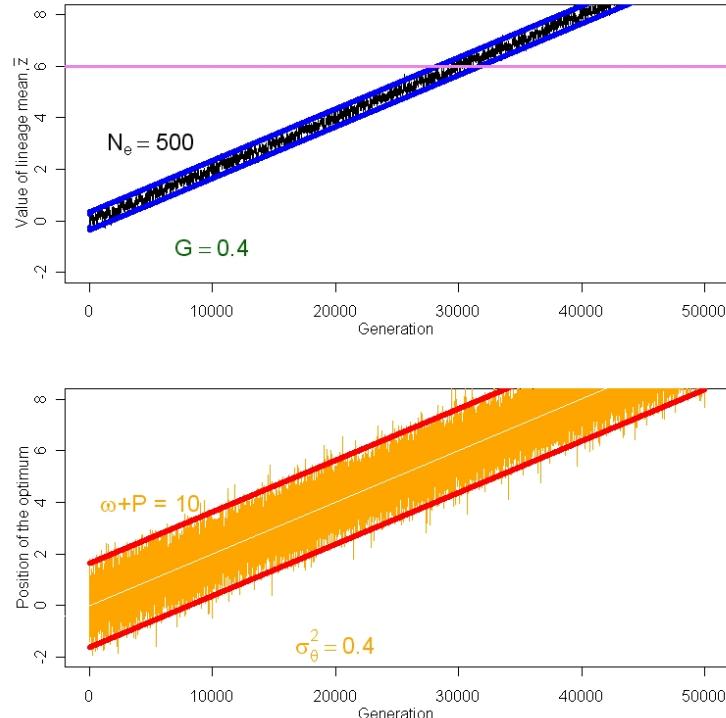
**Figure 13.7** Simulations of 20 replicate populations evolving in response to a very slowly moving intermediate optimum. The rate of motion of the optimum is  $k=0.0002$ , 25 times slower than in Fig. 13.6.

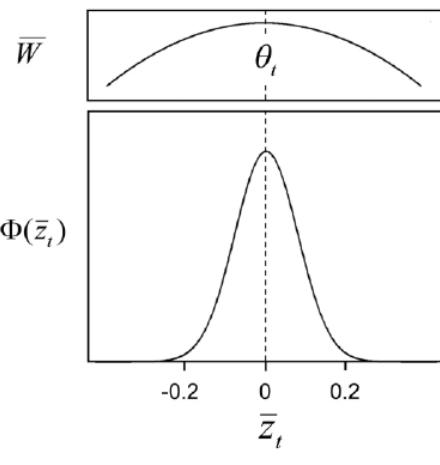
Displacement of the optimum will cause a lag, a temporary exaggeration of distance between mean and the adaptive peak. If the optimum moves at a steady rate, this lag equilibrates at a value that is larger when stabilizing selection is weak and genetic variance is small,

$$L \equiv \theta_t - \bar{z}_t = (\omega + P)G^{-1}\Delta\theta,$$

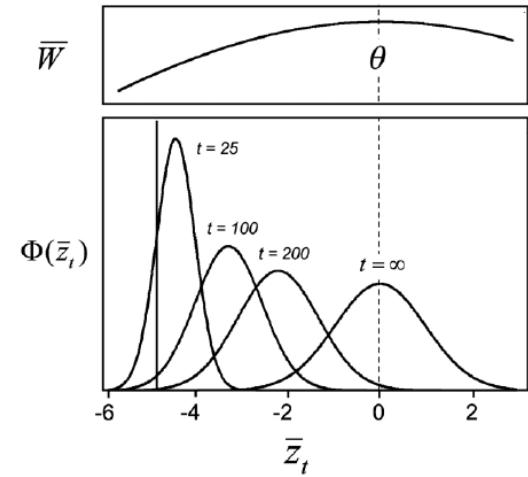
where  $\Delta\theta$  is the per generation change in the position of the optimum

(Lynch & Lande 1993, Jones et al. 2004). Lag can have deleterious demographic consequences for the population. The further the mean is displaced from the adaptive peak, the greater the decline in the population's mean fitness. The resulting impairment in population growth rate can lead to population extinction. Gomulkiewicz & Houle (2009) show how the resulting risk to the population can be expressed in critical values of genetic parameters.





**Figure 4:** Diversification among replicate populations at generation  $t$  under the moving optimum model. The top panel shows the expected position of the optimum,  $\theta_t$ . The bottom panel shows the distribution of phenotypic means,  $\Phi(\bar{z}_t)$ . The expected phenotypic mean lags so very slightly behind the optimum at generation  $t$  that it appears to be superimposed on the optimum. Heritability is 0.4,  $\omega^2 = 10$ ,  $N_e = 1,000$ ,  $k = 0.001$ , and  $\sigma_\theta^2 = 0.001$ . Other conventions as in figure 2.



**Figure 3:** Divergence in response to a displaced optimum. The top panel shows an adaptive landscape with an optimum that has been displaced 5 phenotypic standard deviations from the phenotypic mean, shown as a vertical line. The bottom panel shows the distribution of phenotypic means,  $\Phi(\bar{z}_t)$ , after different intervals of time have elapsed, as populations evolve in response to the displaced optimum. Heritability is 0.4,  $\omega^2 = 99$ , and  $N_e = 50$ . Other conventions as in figure 2.

### 13.3 A single displacement of the adaptive peak

... drift-stabilizing selection balance about a stationary peak (section 11.z) is a special case of this model

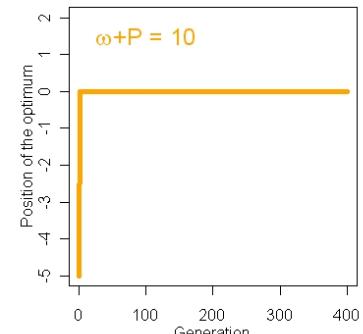
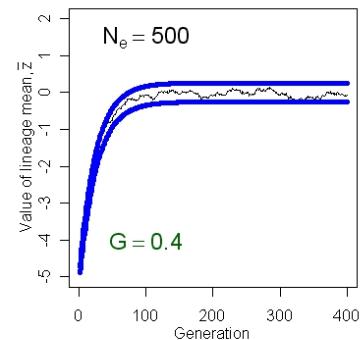
$$E(\bar{z}_t) = \bar{z}_0 \exp\left[-\left(\frac{G}{\omega + P}\right)t\right]$$

$$\text{Var}(\bar{z}_t) = \frac{\omega + P}{2N_e} \left\{ 1 - \exp\left[-2\left(\frac{G}{\omega + P}\right)t\right] \right\}$$

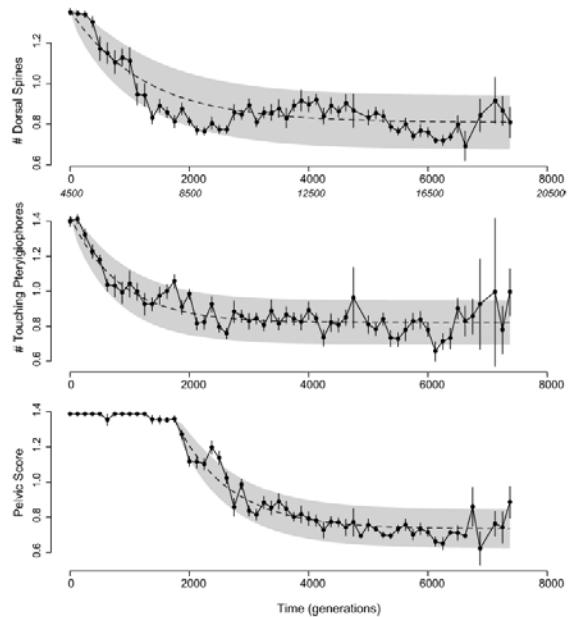
$$\text{Var}(\bar{z}_\infty) = \frac{\omega + P}{2N_e}$$

... (Lande 1976) (Fig. 13.3).

**Figure 13.8.** The trait mean of a single population evolving in response to a single displacement of its adaptive peak. The displacement of the adaptive peak by  $5\sqrt{P}$  occurs instantaneously in generation 1 (lower panel). The evolving trait mean is shown in black in the upper panel. 99% confidence limits under these parameter values are shown in blue.



{Fig. 13.4 = based on Figure 1 from Hunt et al. 2007} ... loss of anti-predator structures in a fossil lineage of stickleback ... good fit to the displaced optimum model ... suggests that an essentially instantaneous escape from predators was followed by a loss of structures over a period of 2,000 generations \*run some simulations; could this really be an instantaneous change in the optimum?\*\*

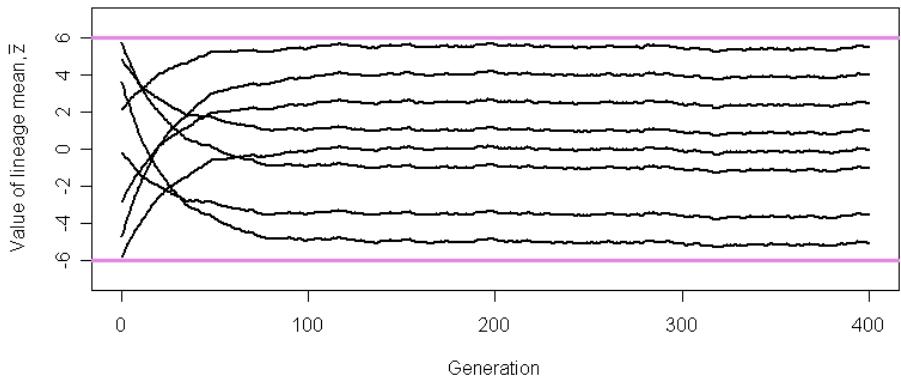


**Figure 1.** Plots of mean trait values over time, starting with the first appearance of the highly armored stickleback lineage. Time is measured in organismal generations (= 2 years/generation), with generation zero set at the first sample marking the invasion of the lake by this particular lineage (smaller italicized numbers below the time axis in the top panel show the time scale from Bell et al. [2006], which is in years). Vertical error bars denote one standard error; samples with fewer than five individuals have been omitted. For each trait, the expected evolutionary trajectory of the best-fit adaptive model is shown as a dotted line, with the 95% probability envelope around this solution in gray. There is a delay before the pelvic score character begins to evolve to a lower armored form (see text).

{add sections here as necessary to cover additional models in the literature ... e.g., the single- and multiple-burst models of Uyeda et al. 2011}

The potential of the displaced optimum model to account for the bounded evolution that is observed in large, long term data sets is illustrated in Fig. 13.10.

**Figure 13.10.** Simulations of 8 lineages evolving in response to different displacements of their intermediate optima. At generation zero, the position of the optima ranges from -5 to 6 and at generation 1 those optima are displaced by  $3.5 \pm 9\sqrt{P}$ .



Here we relax the assumption that the optimum is zero before displacement and in addition allow displacements of different magnitudes in different lineages. These adjustments to the model allow evolution outcomes that produce a band of evolving means that reside of indefinitely within specified boundaries, matching a pattern seen in long term data sets (Estes & Arnold 2007). Nevertheless, even with these adjustments, the model is unsatisfactory in a few respects. Allowing optima displacements only in generation one is obviously artificial. Secondly, we would like both the timing and magnitude of displacements to be generated by a stochastic process, rather than impose them by fiat. The next model grants both of those wishes.

### 13.4 Single burst model

This model corrects the deficiencies of the displaced optimum model with the aim of producing the kind of evolutionary pattern observed in Fig. xx.x (Uyeda et al. 2011). For simplicity in this and the next model, we will equate the position of the trait mean with the position of its intermediate optimum. In other words, we will introduce additional realism into the movement of the optimum, but we will assume that the trait mean track those movements with no lag. We will consider the consequences of this modeling short cut at the end of this section.

A key innovation of this model is that, instead of restricting occurrence of displacements at the beginning of time intervals, we will incorporate stochastic variation in the timing of displacements. The timing of the displacements is generated at random by a Poisson process, but as in the displaced optimum model, only a single displacement is allowed (Fig. 13.14).

**Figure 13.14** Simulation of the single burst model. The orange line shows the position of the optimum. Parameter values are  $\sigma_p^2 = 5$ ,  $\sigma_d^2 = 37$ , and  $\lambda = 0.005$ . The average waiting time for a displacement of the optimum is  $1/\lambda = 200$  generations. The 99% confidence limits for a normal distribution with mean zero and variance  $\sigma_p^2 = 5$  are shown in violet. In this run, a single displacement of the optimum occurs at about generation 770.

A Poisson process is one in which rare occurrences of a certain event, e.g., displacement, are both rare and independent of time. Thus, for a single lineage the probability that no displacement has occurred in elapsed time  $t$  is  $\exp(-\lambda t)$ , where  $\lambda$ , the average rate of occurrence, is a constant. The probability that a first displacement occurs by time  $t$  is  $1 - \exp(-\lambda t)$ . Once a displacement happens, its magnitude,  $d$ , is determined by a draw from a normal distribution with a mean of zero and a variance of  $\sigma_d^2$ . Even in the absence of a displacement, the position of the optimum varies from generation to generation, so that the steady state variance among trait means is normally distributed with a mean of zero and a variance of  $\sigma_p^2$ . The resulting distribution of trait means for a set of replicate lineages is the sum of two normal distributions, each weighted the probability that a displacement does or does not occur, and is given by

$$\Phi(\bar{z}_t) = \frac{[1 - \exp(-\lambda t)]}{\sqrt{2\pi(\sigma_p^2 + \sigma_d^2)}} \exp\left(\frac{-\bar{z}^2}{2(\sigma_p^2 + \sigma_d^2)}\right) + \frac{[\exp(-\lambda t)]}{\sqrt{2\pi\sigma_p^2}} \exp\left(\frac{-\bar{z}^2}{2\sigma_p^2}\right)$$

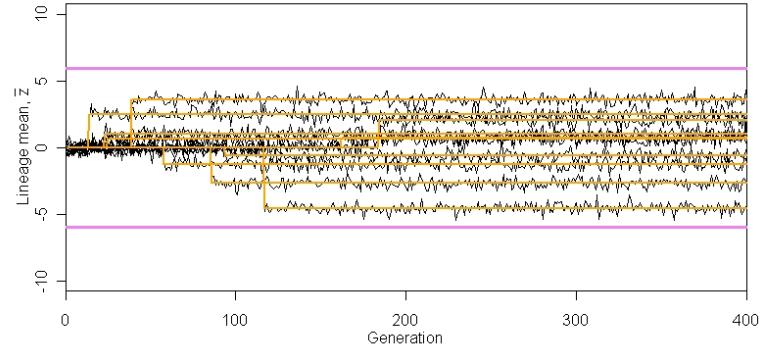
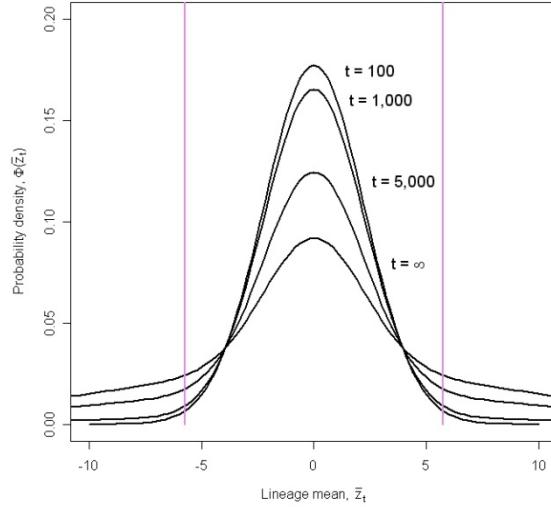
(Uyeda et al. 2011). As  $t \rightarrow \infty$ , this distribution converges on a normal distribution with mean zero and a variance of  $\sigma_p^2 + \sigma_d^2$ .

**Figure 13.11** Distributions of lineage means after various elapsed times according to the single burst model. Parameter values are  $\sigma_p^2 = 5$ ,  $\sigma_d^2 = 87.5$ , and  $\lambda = 0.0001$ . The average waiting time for a displacement of the optimum is  $1/\lambda = 10,000$  generations. The 99% confidence limits for a normal distribution with mean zero and variance  $\sigma_p^2 = 5$  are shown in violet.

When the rate of occurrence,  $\lambda$ , is very small, lineages evolve within a bounded region for thousands of generations. The width of that bounded region is determined by the confidence limits of a normal distribution with a mean at zero and a variance of  $\sigma_p^2$  (Fig. 13.11). On a longer time frame, the bounded region expands so that its width corresponds to the confidence limits of a normal distribution with a mean at zero and a variance of  $\sigma_p^2 + \sigma_d^2$ .

**Figure 13.15.** Simulation of 10 replicate lineages evolving according to the single burst model. The orange lines show the position of the optima, which undergo a single displacement in each lineage.

Parameter values are  $\sigma_p^2 = 0.1$ ,  $\sigma_d^2 = 10$ , and  $\lambda = 0.01$ . The average waiting time for a displacement of the optimum is  $1/\lambda = 100$  generations. Typical bounds for actual data are shown in violet.



In the preceding section we used the single burst model to explain departures from stasis that persists on a long time scale (for hundreds of thousands to millions of generations) but the model can also be used to represent evolution on a much shorter time scales. For example, just as we used the displaced optimum model to explain diversification within but not outside the normal bounds of data (Fig. 13.10), the single burst model can also be employed to the same end. In Fig. 13.15 we have chosen a small value for  $\sigma_p^2$  to represent small amplitude fluctuations in the position of the optimum, and a modest value for

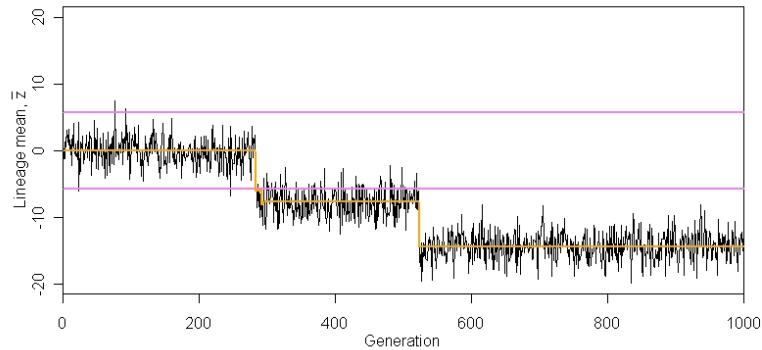
$\sigma_d^2$  so that single displacements of the optimum are unlikely to carry the trait mean more than about 5 standard deviations away from the clade's long term average.

Although this model makes important and useful innovations in how the optimum moves, it takes some short cuts in modeling the evolution of the trait mean. In the first place, we have simply equated the movement of the trait mean with the movement of the optimum. This economy in modeling is obviously a simplification, because we know from earlier models (e.g., section 13.3) that the trait mean is likely to take dozens of generations to track a sudden, large displacement of the optimum. Secondly, we have modeled relatively large displacements of the optimum, but ignored the possibility of fluctuations in position on short time scales. Instead, we employed a white noise process to generate short term stochastic variation in the trait mean via the variance parameter  $\sigma_p^2$ . Again, we know that this is an unrealistic simplifying assumption, because we discovered in sections 13.1 and 13.2 that even white noise movement of the optimum does not evoke white noise movement of the trait mean. In a white noise process, the deviation that is drawn from a normal distribution is from the long term average, not just from the trait mean in the last generation (as in Brownian motion). The consequence is that under a white noise process, the mean would be rapidly whipped from one position to the next every generation. The normal rules of quantitative inheritance prevent this from happening in the real world. A real mean meanders through time (Fig. 13.5a), it does not whip back and forth. In other words,  $\sigma_p^2$  should be taken as a surrogate of variation produced by some process that needs to be more realistically modeled. White noise is an assumption of convenience that produces a desired result. The desired result is that variance produced by the process does not increase with time, enabling us to account for the long term band of stasis that is observed in long term data sets (Fig. 7.x). The upshot is that while the single burst model produces some key features that characterize data (long term stasis or long delays in the appearance of substantial diversification), the model is still a work in process that awaits realistic modeling for short term behavior of the mean. The same realization applies to the model that we shall consider next.

### 13.5 Multiple bursts model

This model corrects another deficiency of the displaced optimum model that was not addressed by the single burst model. The present model allows multiple displacements of the optimum. ‘Multiple bursts’ refers to this key feature of the model. Multiple displacements of an intermediate optimum can occur at random during the lifespan of a lineage. Each of these displacements, may evoke a burst of evolution, but these bursts are interspersed by potentially long intervals during which the optimum is stationary and evolution is static (Fig. 13.13). As in the single burst model, the occurrence of the displacements is generated

**Figure 13.13.** Simulation of the multiple burst model. The orange line shows the position of the optimum. Parameter values are  $\sigma_p^2 = 5$ ,  $\sigma_d^2 = 37$ , and  $\lambda = 0.005$ . The average waiting time for a displacement of the optimum is  $1/\lambda = 200$  generations. The 99% confidence limits for a normal distribution with mean zero and variance  $\sigma_p^2 = 5$  are shown in violet. In this example, three displacements (at about 290, 300, 520 generations) move the lineage mean in the same direction, but that will not be the case in most runs.



generated at random intervals by a Poisson process. The general characteristic of the process is that the probability of observing  $n$  displacement events in some time interval of length  $t$  is given by

$$p_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!},$$

where  $\lambda$  is a constant, the long term average rate of displacement (Bailey 1964). The expected number of displacements in an interval of length  $t$  is  $\lambda t$ . Once a displacement occurs, its magnitude,  $d$ , is determined by a draw from a normal distribution with a mean of zero and a variance of  $\sigma_d^2$ . Let the trait mean be zero at the start of a time interval of length  $t$ . At the end of that interval the probability distribution for the magnitude of divergence,  $\bar{z}$ , is a compound distribution given by

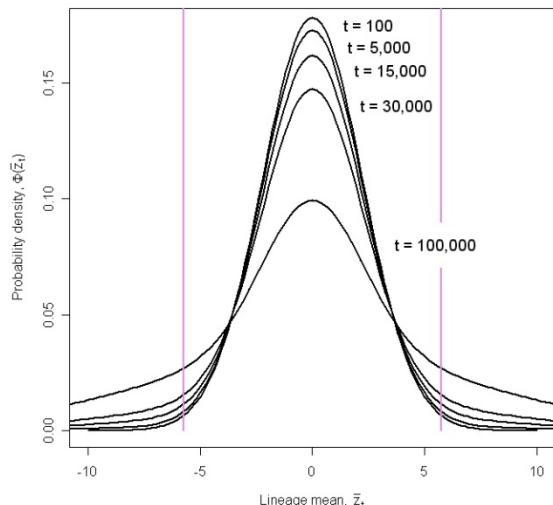
$$\Phi(\bar{z}_t) = \sum_{n=0}^{\infty} \frac{\exp\left(\frac{\bar{z}^2}{2n\sigma_d^2}\right)}{\sqrt{2\pi n\sigma_d^2}} e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

(Uyeda et al. 2011). As before, we add an element of white noise variation,  $\sigma_p^2$ , to the trait mean that represents, for example, variance arising from the balance between drift and stabilizing selection, as well as contributions from other sources. Our probability distribution becomes

$$\Phi(\bar{z}_t) = \sum_{n=0}^{\infty} \frac{\exp\left(\frac{\bar{z}^2}{2(\sigma_p^2 + n\sigma_d^2)}\right)}{\sqrt{2\pi(\sigma_p^2 + n\sigma_d^2)}} e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

(Uyeda et al. 2011). The behavior of this model differs from the single burst model in an important way. Instead of the distribution converging on a limiting variance, the variance expands perpetually. Nevertheless, this model can capture the essential features of the data summarized in Fig. 7.x, if displacements are extremely rare. Under that condition, lineage means can reside for tens thousands of generations within boundaries approximated by the confidence limits of a normal distribution with

variance  $\sigma_p^2$ . On longer time scales, lineage means can diverge beyond those limits. Figure 13.12 shows the distributions of lineage means under the same parameters as in Fig. 13.12, but here the average waiting time for displacement of the optimum is much longer, 100,000 generations, rather than 200. If displacements of the optimum are appreciably rarer than in Fig. 13.12, a initial period of relative stasis can last hundreds of thousands or millions of generations, followed by a period of more substantial divergence that can closely approximate the pattern in Fig. 7.x.



**Figure 13.12** Distributions of lineage means after three different elapsed times according to the multiple burst model. Parameter values are  $\sigma_p^2 = 5$ ,  $\sigma_d^2 = 37$ , and  $\lambda = 0.00001$ . The average waiting time for a displacement of the optimum is  $1/\lambda = 100,000$  generations. The 99% confidence limits for a normal distribution with mean zero and variance  $\sigma_p^2 = 5$  are shown in violet.

Uyeda et al. (2011) used maximum likelihood to fit this model to the data shown in Fig. 7.x and estimated that  $1/\hat{\lambda} = 10^{7.3976}$ . In other words, the average waiting time to a displacement of the optimum was nearly 25 million years, the average magnitude of displacement was only xxx \*need value on standard scale\*. Thus, displacements of the optimum appear to be very rare indeed. Substantial divergence of the magnitude observed in time intervals longer than 10 million years, apparently arise as consequence of repeated instances of rare movements of the optimum, each displacement being of modest magnitude. Substantial macroevolutionary divergence arises from the accumulation of repeated, rare, modest bursts of evolution.

The multiple bursts model was the best performing of three models compared by Uyeda et al. 2011 ... \* more discussion here\*

Before leaving the multiple burst model, we need to recall that it shares certain simplifying features with its more constrained sibling, the single burst model. Both of these models employ a white noise process because of the property that its contribution of among-lineage variance will be time invariant, even though its characterization of short term behavior of the mean is unrealistic. Secondly, as before, we have assumed that the response of the trait mean to the displaced optimum is instantaneous rather than asymptotic. Both of these features need to be replaced by more realistic assumptions, which will inevitably mean more complicated probability expressions to replace 13.x and 13.y \*equations for the two prob distributions\*.

### 13.5 Synapomorphy resulting from a rare displacement of an adaptive peak

A character that changes just once in the phylogenetic history of a clade constitutes an ideal indicator of phylogenetic relationships and is known as a synapomorphy (Hennig 1966) ... The most plausible model of the process that underlies synapomorphy is the one discussed in 13.3, a single displacement of an adaptive peak ... rapidity of adjustment in phenotypic mean ... the most puzzling aspect of synapomorphic peak movement is why the displacement should occur just once ... {direct reader to ecological discussion in Chap 19} ... But whatever the causes of synapomorphic peak movement, the phenomenon is sufficiently common that systematists are routinely able to find characters that adhere to this mode of evolution.

### 13.6 Testing models of peak movement

... {a summary of the Estes & Arnold 2007 results using the Gingerich data = Fig. 13.5 = based on summary figure in Estes & Arnold 2007}

### 13.7 Double OU process

... {present and discuss here depending on what is published by Schweder et al}

### 13.8 Other peak controller processes

## 13.0 Evidence for fluctuation and trends in the position of intermediate optima

A variety of sources suggest that even though the movement of the adaptive peak is conservative enough to be responsible for stasis in the fossil record, small scale movements can be detected in ecological

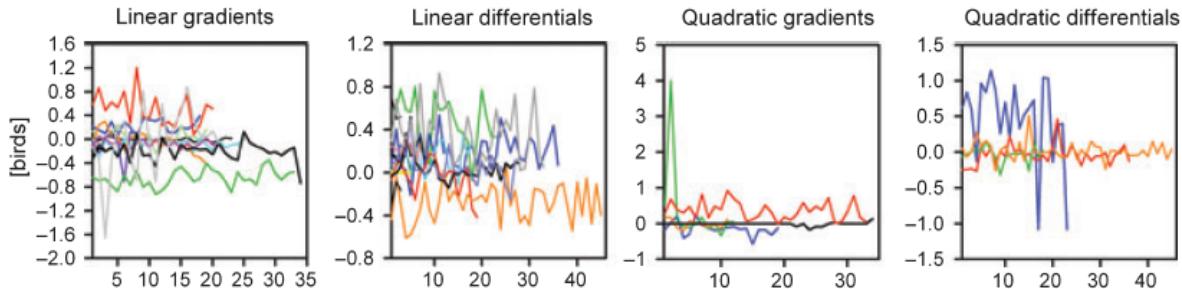
studies. At the finest temporal scale, we can estimate the speed of movement from studies that record trait means generation-to-generation, on the basis of either of two suppositions: that means closely tracks the optimum or that lag has stabilized so that the position of the mean mirrors the position of the optimum. In a sample of 4 such studies (Table 13.x), trait means moved an average of 0.09 phenotypic standard deviations per generation over an average of 11 generations. The average per generation variance in movement of the optimum,  $\sigma_\theta^2$ , estimated from the same studies, was 0.10.

**Table 13.x.** Estimates of the average rate ( $\Delta\theta$ ) and sample variance ( $\sigma_\theta^2$ ) of per-generation shifts in the selective optimum, in units of phenotypic standard deviation. The number of change-time intervals in each study is shown in the  $n$  column. Absolute values of average rates are presented in the  $\Delta\theta$  column. From Jones et al. 2011.

Species	Trait <sup>a</sup>	Number of generations	$n$	Average rate ( $\Delta\theta$ )	Variance ( $\sigma_\theta^2$ )	Ref
Grayling ( <i>Thymallus thymallus</i> )	Age-specific body length	13	6	0.184	0.582	Haugen and Vøllestad (2001), Fig. 3
	at five different ages	13	7	0.092	0.118	
		13	6	0.010	0.036	
		13	6	0.021	0.060	
		13	6	0.045	0.079	
Darwin's finch ( <i>Geospiza fortis</i> )	Body size	6	24	0.068	0.067	Grant and Grant (2002), Fig. 1
	Beak size	6	24	0.000	0.061	
	Beak shape	6	25	0.078	0.040	
Darwin's finch ( <i>Geospiza scandens</i> )	Body size	5	24	0.197	0.062	
	Beak size	5	25	0.137	0.100	
	Beak shape	5	24	0.290	0.025	
Great tit ( <i>Parus major</i> )	Fledgling mass	19	34	0.003	0.0003	Garant et al. (2004), Fig. 3
<b>Mean (SE)</b>		10.8		0.094 (0.026)	0.103 (0.045)	

<sup>a</sup>Phenotypic means were used for all estimates except the great tit dataset, where breeding values were used.

Another perspective is provided by studies that estimate selection coefficients from the same population over a period of years. A sample of time series from studies of bird populations,



**Figure 13.y.** Estimates of selection coefficients year-by-year in long term studies of bird populations. Estimates of 4 different coefficients are plotted as a function of consecutive year in the study. Time series for 6 different studies are plotted in different colors. Quadratic coefficients were estimated in only 4 studies (Siepielski et al. 2009).

among the best studied taxa in this regard, is shown in Fig. 13.y. Fluctuation in selection coefficients is a conspicuous feature of these and other time series compiled by Siepielski et al. 2009. In some case, the fluctuations in directional selection ( $\beta$ ) are about a long term trend of zero. This pattern is consistent with fluctuation of an adaptive peak ( $\theta$ ) about a stationary position ( $\bar{\theta} = 0$ ). In other cases, the fluctuations are about a more less constant value of  $\beta$ , suggesting that the peak is moving at a constant rate, with the trait mean lagging behind that moving optimum. Estimates of the variation in  $\beta$ , which include contributions from both temporal fluctuation in direction and strength (Siepielski et al. 2009, Fig. 4a), suggest a overall median value of about  $\sigma_\beta^2 = 0.01$ , with most values falling in the range 0-0.25. These same studies also indicate that the configuration (shape and orientation) of the AL fluctuates through time. Although the available data are meager and difficult to interpret, they suggest that fluctuations of this sort are modest compared to fluctuations and trends in  $\theta$ . We shall refer to these results in the sections that follow, for they provide a means of validating choices of parameter values in modeling and simulations studies **\*\*need to systematically do that cross-checking\*\***.

The challenge that faces us is to find a model for peak movement capable of accounting for the ecological results that we have just reviewed, but that can also account for the evidence of stasis and evolutionary bursts that springs from data on a paleontological time scale. In the sections that follow, we review a set progressively more complicated models that takes to this goal.