

Neutral Drift of traits

Bruce Walsh lecture notes
Uppsala EQG 2012 course
version 4 Feb 2012

Detailed reading: LW Chapter 11

Outline

- Inbreeding and drift
- Divergence in mean via inbreeding
- Effective population size
- Confidence intervals and tests for variances
- Short-term divergence tests
 - Is pattern consistent with drift?
- Long-term divergence (mutation-driven)
- Long-term divergence tests

Inbreeding and Drift

Inbreeding is just a special case of drift.

Recall that the inbreeding coefficient $f = \text{Prob}(\text{both alleles in an individual are IBD})$

Under Wright-Fisher model, an individual is formed by choosing two random gametes

Draw the first allele, there is a $1/(2N)$ chance that we redraw this same allele

Otherwise, allele is not from the same parent, so that chance it is IBD is just $f(t)$

Putting these together gives the equation
For the increase in f under drift as

$$f(t) = 1/(2N) + [1 - 1/(2N)] f(t-1)$$

Rearranging, this becomes

$$1 - f(t) = [1 - 1/(2N)] * [1 - f(t-1)]$$

Which has solution

$$1 - f(t) = [1 - 1/(2N)]^t * [1 - f(0)]$$

Starting with an initial non-inbred population
($f(0) = 0$), we have

$$f(t) = 1 - [1 - 1/(2N)]^t$$

Divergence based on existing variation

- Recall that a set of initially identical populations under inbreeding
 - Loose within-pop variation $(1-f)\text{Var}(A)$
 - Accumulate a between-population variation (divergence in means) $2f\text{Var}(A)$
- Drift is a case of inbreeding, with
 - $f(t) = 1 - [1 - 1/(2N)]^t \sim 1 - \exp(-t/2N)$
 - Last step follows since $(1-x) \sim 1 - x + O(x^2)$
- Hence, expected divergence based on existing variation is
 - $\sigma^2(\mu) \sim [1 - \exp(-t/2N)] \text{Var}(A)$

What generates the btw-pop variance?

Let the rv X denote the final contribution from a locus.

$$X = \begin{cases} 2a & \text{with probability } p_0 \\ 0 & \text{with probability } 1 - p_0 \end{cases}$$

The between-line variance generated by fixation of the initial additive variance is

$$\begin{aligned} \sigma^2(X) &= E(x^2) - [E(x)]^2 = (2a)^2 * p_0 - (2a p_0)^2 \\ &= 4a^2 * p_0(1 - p_0) = 2 \sigma_A^2(0) \end{aligned}$$

Hence, the between-line variance approaches twice the initial additive variance.

Effective Population size, N_e

When the population is not ideal (changes over time, unequal sex ratio, uneven contribution from individuals), we can still compute an effective population size N_e which gives the size of an ideal population that behaves the same as our population

We will consider N_e under

- population bottlenecks
- unequal sex ratio
- unequal contribution for all individuals

N_e under varying population size

If the actual population size varies over time, the effective population size is highly skewed towards the smallest value

If the populations sizes have been $N(1), N(2), \dots, N(k)$, the effective population size is given by the harmonic mean

$$N_e = \frac{k}{\sum_{i=1}^k \frac{1}{N(i)}}$$

Suppose the population sizes are 10000, 10000, 10000, 100.

N_e becomes 399

If 10000 is replaced by 10^9 , N_e becomes 400

N_e under unequal sex ratios

When there are different number of males (N_m) and females (N_f), the effective population size is skewed towards the rarer sex

$$N_e = \frac{4N_m * N_f}{N_m + N_f}$$

For example, suppose we used 2 male salmon to fertilize the eggs of 1000 females. What is N_e in this case?

$$N_e = (4*2*1000)/(2 + 1000) = 8$$

N_e under unequal individual contributions

Not all individuals contribute equally to the next generation. What effect does this have on N_e ?

Let σ^2_0 be the variance in offspring number for individuals in the population, then

$$N_e \simeq \frac{2N}{\sigma^2_0/2 + 1}$$

If contributions follow a Poisson with a mean of 2 offspring per parent (male + female replace each other), then $\sigma^2_0 = 2$, and $N_e = N$

N_e under unequal individual contributions (cont)

$$N_e \simeq \frac{2N}{\sigma_O^2/2 + 1}$$

If all individuals contribute EXACTLY the same number of offspring, $\sigma_O^2 = 0$, and $N_e = 2N$, so that the effective pop size is twice the actual size

In a survey of reproductive success in birds, Grant found that $\sigma_O^2/2$ ranged from 1.2 to 4.2, giving an N_e of only 40 - 90% of the actual number of females

Summary: short-term divergence

- If we start n replicate populations, then after t generation the variance between them (the variance of their means), $\text{Var}(B)$ has expected value
 - $\text{Var}(B) \sim [1 - \exp(-t/2N_e)] \text{Var}(A)$
- Hence we can ask, given N_e , of whether an observed divergence is
 - Too large (direction selection)
 - Too small (stabilizing selection)
- First, we require some distribution theory on variance estimators

Suppose $y_i \sim N(\mu, \sigma^2)$, Let $\text{Var}(y) = [n-1]^{-1} \Sigma (y_i - \bar{y})$

Then $(n-1)\text{Var} \sim \sigma^2 * \chi^2_{n-1}$

(i) Confidence intervals for σ^2

As a result, confidence intervals for the true variance σ^2 based on the observed sample variance Var follow from critical values for a χ^2 distribution. Letting $X_{p,n}$ satisfy $\Pr(\chi_n^2 \leq X_{p,n}) = p$, then

$$\Pr(X_{\alpha/2,n} \leq \chi_n^2 \leq X_{1-\alpha/2,n}) = 1 - \alpha \quad (11.5b)$$

From Equation 11.5a, substituting $(n-1)\text{Var}/\sigma^2$ for χ_{n-1}^2 , we have

$$\Pr\left(X_{\alpha/2,n-1} \leq \frac{(n-1)\text{Var}}{\sigma^2} \leq X_{1-\alpha/2,n-1}\right) \quad (11.5c)$$

$$= \Pr\left(\frac{1}{X_{\alpha/2,n-1}} \geq \frac{\sigma^2}{(n-1)\text{Var}} \geq \frac{1}{X_{1-\alpha/2,n-1}}\right) = 1 - \alpha \quad (11.5d)$$

giving

$$\Pr\left[\left(\frac{n-1}{X_{1-\alpha/2,n-1}}\right) \text{Var} \leq \sigma^2 \leq \left(\frac{n-1}{X_{\alpha/2,n-1}}\right) \text{Var}\right] = 1 - \alpha \quad (11.6)$$

(ii) Significant testing

We can also use Equation 11.5c to assess the significance of an observed sample variance given some assumed value σ_0^2 . Rearranging Equation 11.5c gives

$$\Pr \left[\left(\frac{\sigma_0^2}{n-1} \right) X_{\alpha/2, n-1} \leq \text{Var} \leq \left(\frac{\sigma_0^2}{n-1} \right) X_{1-\alpha/2, n-1} \right] = 1 - \alpha \quad (11.7)$$

An observed sample variance outside of this interval is said significantly different at level α from that expected under the null.

(iii) Power

If the true variance is really $\sigma_1^2 \neq \sigma_0^2$, then the **power** (LW Appendix 5) for this parameter value is just the probability that a sample variance falls outside of the interval given by Equation 11.7, which is a function of the sample size n and the assigned significance α for the test. Letting β denote the probability of a type II error (failing to declare a test significant when the null is false), we can obtain this from Equation 11.7 by noting that now $[(n-1)/\sigma_1^2] \text{Var} \sim \chi_{n-1}^2$. Multiplying all terms of Equation 11.7 by $(n-1)/\sigma_1^2$ gives the probability β of a sample variance failing to be declared significant as

$$\begin{aligned}\beta &= \Pr \left[\left(\frac{\sigma_0^2}{n-1} \right) \left(\frac{n-1}{\sigma_1^2} \right) X_{\alpha/2, n-1} \leq \chi_{n-1}^2 \leq \left(\frac{\sigma_0^2}{n-1} \right) \left(\frac{n-1}{\sigma_1^2} \right) X_{1-\alpha/2, n-1} \right] \\ &= \Pr \left[\left(\frac{\sigma_0^2}{\sigma_1^2} \right) X_{\alpha/2, n-1} \leq \chi_{n-1}^2 \leq \left(\frac{\sigma_0^2}{\sigma_1^2} \right) X_{1-\alpha/2, n-1} \right] \quad (11.8a)\end{aligned}$$

Hence, the power $1 - \beta$ is

$$\Pr \left[\chi_{n-1}^2 \leq \left(\frac{\sigma_0^2}{\sigma_1^2} \right) X_{\alpha/2, n-1} \right] + \Pr \left[\chi_{n-1}^2 \geq \left(\frac{\sigma_0^2}{\sigma_1^2} \right) X_{1-\alpha/2, n-1} \right] \quad (11.8b)$$

Example 11.1. Consider a sample variance estimated from $n = 10$ observations (e.g., the between-group variance estimated from the means of ten replicate lines). Since $\Pr(\chi_9^2 \leq 2.700) = 0.025$ and $\Pr(\chi_9^2 \leq 19.023) = 0.975$, Equation 11.6 gives the 95% confidence interval ($\alpha = 0.05$) on the true variance σ^2 as between $(9/19.023)\text{Var}$ and $(9/2.7)\text{Var}$, or $0.473 \cdot \text{Var}$ to $3.333 \cdot \text{Var}$, for an uncertainty in σ^2 spanning almost a full order of magnitude.

What observed values of the sample variance are unlikely given an assumed variance of σ_0^2 ? From Equation 11.7, the upper and lower critical values (for a two-sided test with $\alpha = 0.05$) are $(2.700/9)\sigma_0^2 = 0.3 \cdot \sigma_0^2$ and $(19.023/9)\sigma_0^2 = 2.11 \cdot \sigma_0^2$. Finally, what is the power of this design (again taking $\alpha = 0.05$) when $\sigma_1^2 = \sigma_0^2/2$? Equation 11.8b gives the power as

$$\Pr\left(\chi_9^2 \leq \frac{2.700}{2}\right) + \Pr\left(\chi_9^2 \geq \frac{19.023}{2}\right) = 0.39$$

and hence a type II error rate of 61% when the true variance is half the assumed variance. A similar calculation assuming $\sigma_1^2 = 2\sigma_0^2$ gives a power of 0.20, or a type II error rate of 80%. Useful **R** commands for these calculations are **pchisq(x, n)**, which returns $\Pr(\chi_n^2 \leq x)$, and hence **1 - pchisq(x, n)** returns $\Pr(\chi_n^2 \geq x)$, while **qchisq(p, n)** returns $X_{p,n}$.

In class problem

- Suppose $n = 30$, $\text{Var} = 50$
- What is a 95% CI for the true variance?
- Is this variance significantly different (at $\alpha = 0.05$) from 30?
- Suppose the true variance is 60. What is the power for a test of $\alpha = 0.05$ that the true variance is 30.

Tests of short-term divergence

- Assume the divergence time is sufficiently short that new mutation can be ignored, so between-line divergence is entirely due to drift partitioning the initial variance.
- This gives rise to Lande's (1977) F test

Lande's F Test

Is an observed divergence over a modest amount of time significantly different than expected by drift? For the case in which one has only a single estimate of the among-population divergence, Lande (1977) suggested the statistic

$$F = \frac{\text{Var}(B, t)}{t \cdot \text{Var}(A, 0)/N_e} \quad (11.9a)$$

as a test for neutrality. As noted by Lande, under approximate assumptions, this follows an F distribution, with $L - 1$ numerator degrees of freedom and ∞ denominator degrees of freedom (see notes).

Example 11.2. Lande (1977) used Equation 11.9a to evaluate the results of a 12-year divergence experiment involving five populations of *Drosophila pseudoobscura* (Anderson 1973). Two of the populations had been maintained at 25°C, two at 27°C, and one at 16°C. They were then raised in two common environments (16 and 25°C) and measured for wing length. Estimates of the additive genetic variance for these two environments were 0.88 and 0.77, while the among-population variances were approximately 6.62 and 4.37 respectively. An approximate upper bound for the number of generations of divergence is $t = 150$, whereas the effective population size probably always exceeded $N_e = 1000$. The use of these extreme bounds gives conservative estimates of F , making it more difficult to demonstrate diversifying selection on wing length. Even so, the ratios of observed to expected among-population variance are 50 and 38, both of which are highly significant (comparing these with the critical F ratio with four degrees of freedom in the numerator, and infinite degrees of freedom in the denominator). Thus, the hypothesis that the observed line divergence is solely attributable to random genetic drift can be rejected confidently. More likely, the different thermal conditions resulted in selection for different wing lengths.

Long-term divergence

- At some point, all the initial variation is fixed between populations, and new mutations start to become important in the divergence process
- Further, these new mutations eventually result in a new equilibrium level of additive variation

Mutation-Drift Equilibrium: Within-population Additive Variance

With quantitative traits, the mutation *rate* is not a sufficient descriptor, as we also need a measure of effect size as well.

The mutational variance σ^2_m is thus the measure of the mutational input for quantitative traits

Typically estimated by rate of divergence between highly inbred lines or selection response in such lines

Typical values are on the order of 1/1000 of the environmental variance, e.g.

$$\sigma_m^2 \sim 10^{-3} \sigma_E^2$$

Dynamics of additive variance under mutation & drift:

$$\sigma_A^2(t+1) = \sigma_m^2 + \left(1 - \frac{1}{2N_e}\right) \sigma_A^2(t)$$

At equilibrium, $\hat{\sigma}_A^2 = 2N_e \sigma_m^2$

The general solution is given by

$$\sigma_A^2(t) \simeq 2N_e\sigma_m^2 + [\sigma_A^2(0) - 2N_e\sigma_m^2] \exp(-t/2N_e)$$

The fraction of the variance at any time due to new mutation is just

$$\sigma_{A,m}^2(t) \simeq 2N_e\sigma_m^2 [1 - \exp(-t/2N_e)]$$

The equilibrium heritability becomes

$$\hat{h}^2 = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2 + \sigma_E^2} = \frac{2N_e\sigma_m^2}{2N_e\sigma_m^2 + \sigma_E^2}$$

We can alternatively express the equilibrium heritability as

$$\hat{h}^2 = \frac{2N_e\sigma_m^{2*}}{2N_e\sigma_m^{2*} + 1}, \quad \text{where} \quad \sigma_m^{2*} = \frac{\sigma_m^2}{\sigma_e^2}$$

Note the strong similarity to the equilibrium heterozygosity under Crow-Kimura

Example: Suppose $\sigma_m^2 \sim 10^{-3} \sigma_E^2$ and $N_E = 2$

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_E^2} = \frac{4 \times 10^{-3} \sigma_E^2}{4 \times 10^{-3} \sigma_E^2 + \sigma_E^2} = 0.004$$

Taking $N_E = 500$ gives $h^2 = 0.5$

Complications/Corrections with H and h^2

Using the typical value of $\sigma_m^2 \sim 10^{-3} \sigma_E^2$

$$\hat{h}^2 = \frac{2N_e}{2N_e + 1000}$$

Hence, even a modest value of N_e gives an equilibrium heritability close to one.

The same is seen in the Crow-Kimura result, as when $4N_e\mu > 1$, essentially 100% heterozygotes

Neither of these are seen in natural populations.
What's the problem?

Solution: μ and σ_m^2 are the neutral fractions of the mutation rate and mutational variance

Hence, under drift, H can be NO LARGER than predicted by Crow-Kimura. Same for h^2 . However, both can be MUCH smaller than predicted from estimates of μ and σ_m^2

If mutants have effects on both the trait and fitness, deleterious ones removed.

Estimates of σ_m^2 typically done in highly inbred lines = small N_e = weak effects of selection against deleterious mutations

Between-line Divergence Under Drift & Mutations

The between-line variance (in generation t) from mutation is

$$\sigma_{B,m}^2(t) = 2\sigma_m^2 [t - 2N_e(1 - e^{-t/2N_e})]$$

The asymptotic rate of divergence is just

$$\sigma_{B,m}^2(t) \simeq 2t\sigma_m^2$$

Note the similarity with single-locus neutral theory, where $d(t) = t\mu$. Both divergence rates are independent of N_e .

Long-term divergence in means

- The long-term divergence in a quantitative trait under drift has been model by two processes
- **Brownian motion**: random motion with no constraints. Let b = per time variance,
 - $\mu(t) \sim N(\mu(0), bt)$
- **Ornstein-Uhlenbeck process**: random motion with with a resorting force pulling back towards zero (a particle under Brownian motion force that is tied by a rubber band to the origin)
 - Model for stabilizing selection

Lande (1976) used a Brownian motion model for a quantitative trait with constant additive variance.

For a strictly additive trait, per-generation sampling variance is $b = \sigma^2_A/N_e$

Hence, distribution of mean at time t ,
 $\mu(t) \sim N(\mu(0), t\sigma^2_A/N_e)$

One measure of drift is how quickly the means diffuse. What is the time until a 50% probability that the means are more than K SDs away from the initial mean?

$$\begin{aligned} \Pr(|x_t - \mu_0| \geq K\sigma_z) &= \Pr\left[\frac{|x_t - \mu_0|}{\sigma_t} \geq \frac{K\sigma_z}{\sigma_t}\right] \\ &= \Pr\left[|U| \geq \frac{K\sigma_z}{\sigma_t}\right] = 0.5 \end{aligned}$$

$$\Pr(|U| \geq 0.675) = 0.5$$

$$\text{Hence, } K\sigma_z/\sigma_t = K\sigma_z/\sigma_A(t/N_e)^{1/2} = 0.675$$

Using $h^2 = \sigma_A^2/\sigma_z^2$ and solving for t gives

$$t = \frac{K^2 N_e}{h^2 0.675^2} \simeq 2 N_e \frac{K^2}{h^2}$$

Example: $N_e = 10$,
 $h^2 = 0.5$. Takes 360
generations for $K = 3$

Constant variance or equilibrium variance?

Lande's original model assumed a constant additive variance, independent of N_e . This gave the between-means variance at time t as $t\sigma_A^2/N_e = th^2\sigma_z^2/N_e$

However, as we have seen the long-term mutation-drift (MD) balance gives $\sigma_A^2 = 2N_e \sigma_m^2$

Hence, $t\sigma_A^2/N_e = t(2N_e \sigma_m^2)/N_e = 2t \sigma_m^2$ is the divergence variance is we use the MD equilibrium value

Recall our previous example for a population
with $N_e = 10$ and $h^2 = 0.5$.

Under constant-(original) variance, it took 360 gen. for
50% of the lines to be 3 SD (or greater) from the mean

Under MD equilibrium, $K\sigma_z/\sigma_t = K\sigma_z/(2t\sigma_m^2) = 0.675$

$$t = K^2 \sigma_z^2 / \sigma_m^2 \sim K^2 (N_e + 1/h_m^2)$$

Typical value for mutational heritability $h_m^2 \sim 0.006$

For above N_e , gives 1680 generations

For $N_e = 100$, $t = 3600$ gens under Lande,
while 3300 gens under MD model

Drift and Divergence in the Fossil Record

Lande's 1976 paper asked if the observed amount of divergence in a fossil sequence is consistent with drift, or if selection has to be invoked.

Assume (first) the constant-variance model,

$$b = \sigma_A^2 / 2N_e$$

$$\mu(t) \sim \text{Normal} \left(\mu(0), \frac{t\sigma_A^2}{N_e} \right)$$

Suppose we observed a divergence of $d = | \mu(t) - \mu(0) |$

$$\begin{aligned} \Pr (| \mu(t) - \mu(0) | \leq d) &= \Pr \left(\frac{ | \mu(t) - \mu(0) | }{ \sqrt{t\sigma_A^2/N_e} } \leq \frac{d}{ \sqrt{t\sigma_A^2/N_e} } \right) \\ &= \Pr \left(|U| \leq \frac{d}{ \sqrt{t\sigma_A^2/N_e} } \right) \end{aligned}$$

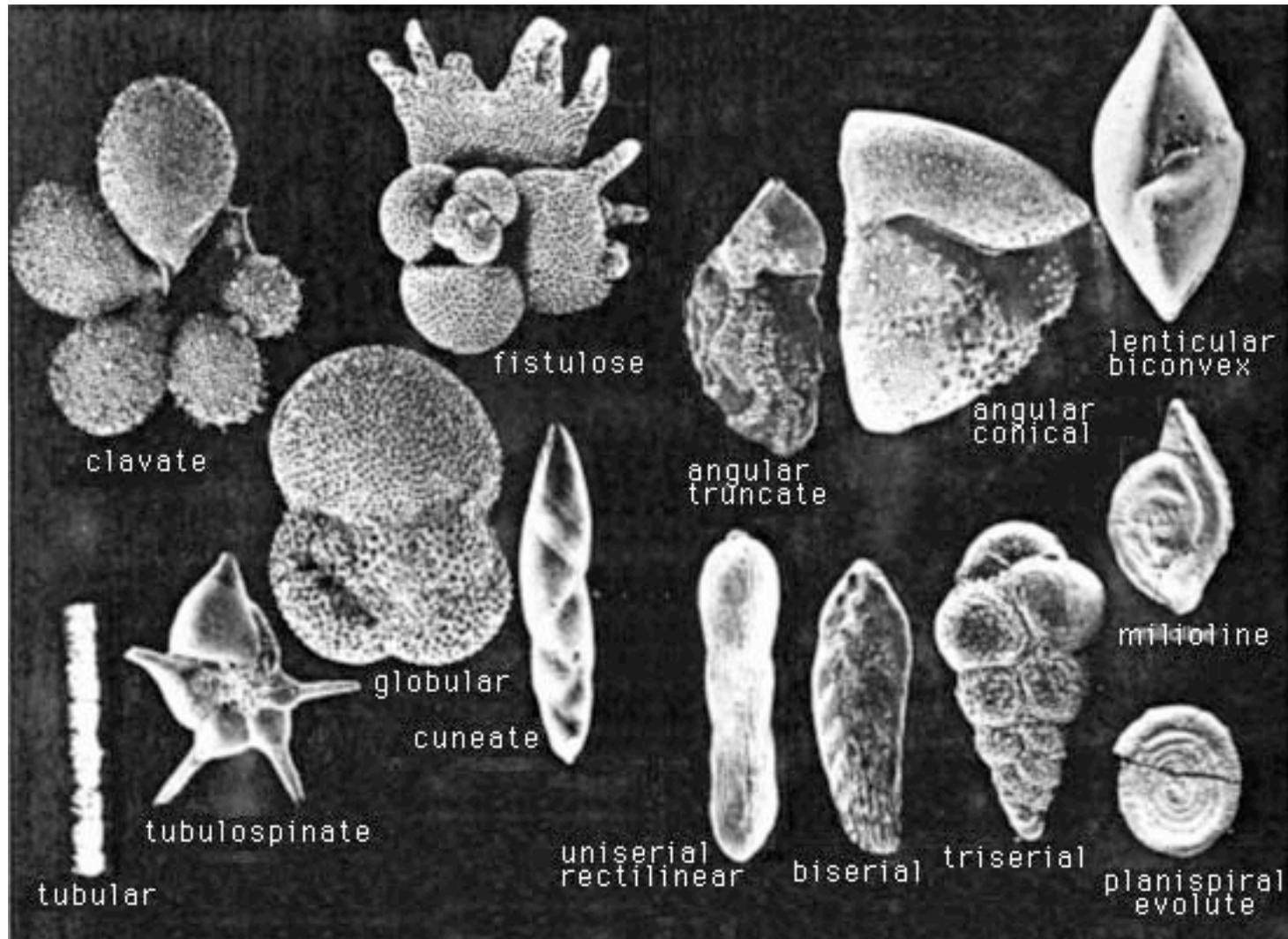
Recalling $\Pr(|U| \leq 1.96) = 0.95$, Lande solved for the largest population size that was consistent with this amount of divergence,

$$1.96 = \frac{d}{\sqrt{t\sigma_A^2/N_e}}$$

Using $\sigma_A^2 = h^2 \sigma_z^2$ and
Solving for N_e gives

$$\widehat{N_e} = \frac{t * h^2 * 1.96^2}{d_*^2} = 3.84 \cdot \frac{t h^2}{d_*^2}$$

foraminifera



Example: Reyment's Foram data

Reyment (1982) observed a 1.49 SD size change over roughly 500,000 generations in a Cretaceous foraminifer.

Taking a typical value of $h^2 = 0.2$,

$$\widehat{N}_e = 3.84 \cdot \frac{t h^2}{d_*^2} = 3.84 \cdot \frac{5 \times 10^5 * 0.3}{1.49^2} \simeq 260,000$$

Paleontological data suggests that the effective population size was greater than 10^6 . Hence, Lande's test suggests that drift could not account for such a rapid divergence.

Turelli et al (1998) offer two important Modifications of Lande's test

First, Lande's test is really two-sided

N_e may be too large to account for the
observed divergence (too fast for drift)

Since $\Pr(|U| > 2.24) = 0.025$, critical value is

$$\widehat{N}_e(\text{fast}) = \frac{t \cdot h^2 \cdot 2.24^2}{d_*^2} = 5.02 \cdot \frac{t h^2}{d_*^2}$$

If N_e exceeds this critical value, we reject
 H_0 that drift can account for rate of change

Likewise, N_e may be too small to account for the observed divergence (too slow for drift).

i.e. consistent with stabilizing selection

Since $\Pr(|U| < 0.031) = 0.025$, critical value is

$$\widehat{N}_e(\text{slow}) = \frac{t \cdot h^2 \cdot 0.031^2}{d_*^2} = 0.00096 \cdot \frac{t h^2}{d_*^2}$$

If N_e is less than this critical value, we reject H_0 that drift can account for rate of change

Second improvement of Turelli et al

Replace σ^2_A with $2N_e \sigma^2_m$. Here

$$\mu(t) \sim \text{Normal}(\mu(0), t\sigma^2_m)$$

Now the test is not based on critical values of N_e , but rather critical values of σ^2_m

We will use the scaled mutational variance,

$$\sigma^2_{m^{**}} = \frac{\sigma_m^2}{\sigma_z^2} = \frac{\sigma_m^2}{(1 - h^2)\sigma_e^2} = \frac{\sigma_m^2 / \sigma_e^2}{1 - h^2}$$

We expect $\sigma^2_{m^{**}}$ values to fall within 10^{-2} to 10^{-4} , although these values can be upper limits if some new mutants also have effects on fitness

The divergence is too fast to be accounted for by drift if the mutational variance is less than the critical value

$$\sigma_{m^{**}}^2 (fast) < \frac{d_*^2}{2t \times 2.24^2} = \frac{d_*^2}{t \times 10.04} = 0.10 \cdot \frac{d_*^2}{t}$$

The divergence is too slow to be accounted for by drift if the mutational variance is greater than the critical value

$$\sigma_{m^{**}}^2 (slow) > \frac{d_*^2}{2t \times 0.031^2} = \frac{d_*^2}{t \times 0.002} = 520.29 \cdot \frac{d_*^2}{t}$$

Let's re-examine Reymont's foram data
Lande's analysis showed that the divergence
was too fast relative to N_e

Under the Turelli et al. Analysis, divergence
is too fast for drift if the mutation variance
is below

$$\sigma_{m^{**}}^2 < 0.10 \times \frac{1.49^2}{5 \times 10^5} = 4.4 \times 10^{-7}$$

This is several orders of magnitude below
typical values and hence drift can easily
account for this divergence.

Bookstein's Modification

Let D_*^2 be the maximal deviation **anywhere** along the time sequence

$$\sigma_{m^{**}}^2 < \frac{D_*^2}{2t \cdot 2.50^2} = \frac{D_*^2}{t \cdot 12.50} = 0.08 \cdot \frac{D_*^2}{t}$$

$$\sigma_{m^{**}}^2 > \frac{D_*^2}{2t \cdot 0.56^2} = \frac{D_*^2}{t \cdot 0.6272} = 1.59 \cdot \frac{D_*^2}{t}$$

Potential problem: Should depend on number of points in the sample.

Ornstein-Uhlenbeck models of Drift

The Ornstein-Uhlenbeck process is Brownian motion with a linear restoring force that tends to move the mean back to zero $\Delta\mu(x) = -ax, \text{Var}(x) = b$

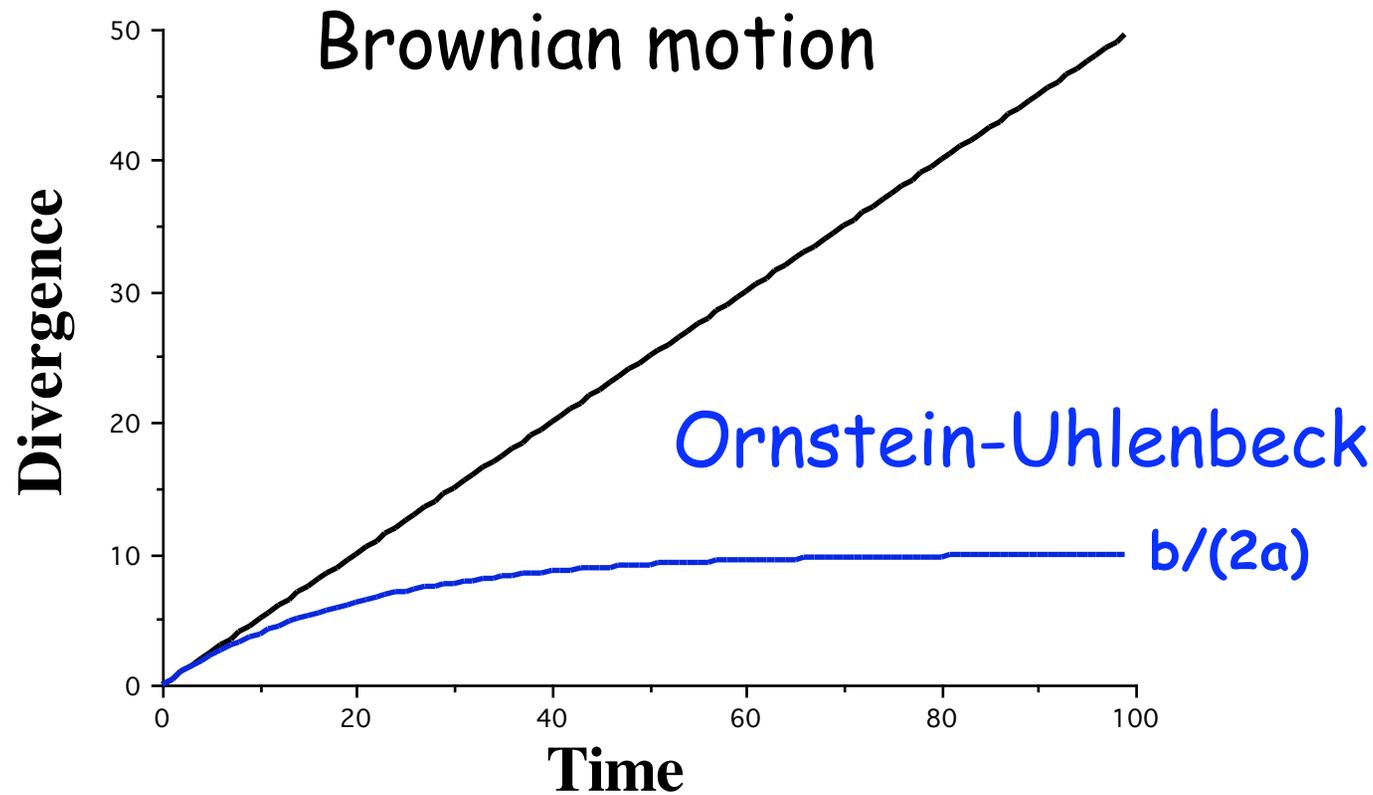
This is a reasonable initial model of drift under stabilizing selection.

The distribution of x_t , given we start at x_0 , is normal, with mean and variance

$$\mu_t = x_0 e^{-at} \quad \sigma_t^2 = \frac{b}{2a} (1 - e^{-2at})$$

Thus for large t , the distribution is $N(0, b/[2a])$

Under OU, divergence reaches a limiting value



Continuous linear divergence under Brownian motion

Divergence approaches a limit under Ornstein-Uhlenbeck

Divergence in gene expression

- Using microarrays (or RNASeq), one can quantify the levels of expression (amount of mRNA) produced by a gene
- For any particular gene, there is both within- and between-population variance in this trait.
- Is that observed pattern of divergence in expression levels more consistent with pure drift, directional selection, or stabilizing selection?

Suggestions of pure drift

- Khaitovich et al -- suggest drift for the divergence in gene expression over the great apes
 - Correlation (over genes) between within- and between-population variation (as might be expected under neutrality)
 - Linear divergence (with generations) between expression levels (as expected under drift with $d_{\dagger} = 2ut$)
 - Problem: Used human probes, and hence divergence of probes results in less binding, and hence more apparent divergence in expression levels

Suggestions of stabilizing selection

- Estimates of the ratio of mutational variance to divergence were too small
 - In *C. elegans*, ratio around 400, = $4N_e$, suggesting an $N_e = 100$ required to account for drift pattern (too small)
 - Similar results for *Drosophila*. Mice, human/chimp

Example 11.5. As developed in Appendix 1, the Ornstein-Uhlenbeck process provides a model of Brownian motion drift coupled with a resorting force back to some optimal value θ , as might be expected with drift and stabilizing selection. Bedford and Hartl (2009) used such a process to fit the pattern of expression divergence within a clade of seven species of *Drosophila*. Under the Ornstein-Uhlenbeck (OU) model, the expected change in the mean value of a process at value x is $a(\theta - x)$, so that if $x < \theta$, it increases, while for $x > \theta$ it decreases. The parameter a which measures the strength of the restoring force is also a measure of the strength of stabilizing selection. As with Brownian motion, the value of the process at time t is normally distributed (Equation A1.33b), but now with mean and variance

$$\mu_t = x_o \exp(-at) + \theta[1 - \exp(-at)], \quad \sigma_t^2 = \frac{b}{2a}[1 - \exp(-2at)]$$

Thus for large t the mean value approaches the optimal value θ while the divergence variance saturates at a value $b/(2a)$, where $b = \sigma_A^2/N_e$ under the constant variance model, giving an asymptotic variance of $\sigma_A^2/(2N_e a)$. Bedford and Hartl found that, in accordance with the OU model, the divergence variance does not linearly increase with time, but rather quickly approaches an asymptotic value. They also introduced a maximum likelihood estimator for a (and hence the strength of stabilizing selection) using divergence data.

F_{ST} vs. Q_{ST}

If trait divergence reflects neutral gene divergence, then F_{ST} should track a similar measure of divergence for traits.

Let Q_{ST} denote the level of population subdivision for allele frequencies at the loci underlying the trait. Letting the genetic variance for the trait in the entire metapopulation under the assumption of panmixia be σ_G^2 , then from the theory developed earlier in this chapter as well as in Chapter 10, the within- and among-subpopulation components of variance can be represented as $\sigma_{GW}^2 = (1 - Q_{ST})\sigma_G^2$ and $\sigma_{GB}^2 = 2Q_{ST}\sigma_G^2$, respectively. It follows that

$$Q_{ST} = \frac{\sigma_{GB}^2}{\sigma_{GB}^2 + 2\sigma_{GW}^2} \quad (11.22)$$

While the term Q_{ST} is due to Spitze (1993), Prout and Barker (1989) and Lande (1992) also proposed this approach.

F_{ST} vs. Q_{ST}

- If neutral divergence, then $F_{ST} \sim Q_{ST}$.
- If adaptation to specific subpopulations, then $F_{ST} < Q_{ST}$.
- If stabilizing selection, then $F_{ST} > Q_{ST}$.
- Caveat: with nonadditive variation, Q_{ST} does not track F_{ST} , with $F_{ST} > Q_{ST}$ expected even for neutral traits.