Lecture 24:
Multivariate Response:
Changes in G

Bruce Walsh lecture notes
Synbreed course
version 10 July 2013

Overview

- Changes in G from disequilibrium (generalized Bulmer Equation)
- Fragility of covariances to allele frequency change
- Resource-partitioning models and genetic correlations
- Long-term directional selection: drift
- Long-term directional selection: mutation
- Long-term quadratic (stabilizing) selection

Infinitesimal Model: Changes in G from linkage disequilibrium

The "Bulmer effect" --- changes in variances from LD produced by selection also applies to covariances

$$\sigma_t(A_i, A_j) = \sigma_0(A_i, A_j) + d_t(i, j)$$

Hence, G(t) = G(0) + D(t), where D is a matrix of all pairwise disequilibrium contributions

Dynamics of D as with univariate case: each generation half the value removed by recombination (for unlinked loci)

3

If d_t is the current LD and d_t* is the LD generated by selection of the parents, then the LD in offspring changes by

$$\Delta d_t(i,j) = -\frac{d_t^*(i,j) + d_t(i,j)}{2}$$

The LD generated by selection is just **G**$_t$* - **G**$_t$, where **G*** is the **G** matrix after selection (but before reproduction)

$$\Delta \mathbf{D}_t = -\frac{\mathbf{G}_t^* - \mathbf{G}_t + \mathbf{D}_t}{2}$$

Where phenotypes and breeding values are multivariate-normally distributed, then the within-generation change in **G** is a function of the within-generation change in **P**,

$$\mathbf{G}^* - \mathbf{G} = \mathbf{G}\mathbf{P}^{-1}(\mathbf{P}^* - \mathbf{P})\mathbf{P}^{-1}\mathbf{G}$$

Putting these together gives the multivariate version of the Bulmer equation (due to Tallias) as

$$\Delta \mathbf{D}_t = \frac{1}{2}\left(\mathbf{G}_t\mathbf{P}_t^{-1}(\mathbf{P}_t^* - \mathbf{P}_t)\mathbf{P}_t^{-1}\mathbf{G}_t - \mathbf{D}_t\right)$$

$$= \frac{1}{2}\left(\mathbf{G}_t\mathbf{P}_t^{-1}\Delta\mathbf{P}_t\mathbf{P}_t^{-1}\mathbf{G}_t - \mathbf{D}_t\right)$$

The quadratic selection gradient, an n x n matrix γ, measures the strength of selection on trait variances and covariances

$$\boldsymbol{\gamma} = \mathbf{P}^{-1}\,\boldsymbol{\sigma}\left[\,w,\,(\mathbf{z}-\boldsymbol{\mu})(\mathbf{z}-\boldsymbol{\mu})^{T}\,\right]\mathbf{P}^{-1} = \mathbf{P}^{-1}\,\mathbf{C}\,\mathbf{P}^{-1}$$

Here, γ_{ij} measures the direct selection on the combination of i and j

- $\gamma_{ii} < 0$. Convex selection on trait i. Selection to decrease variance

- $\gamma_{ii} > 0$. Concave selection on trait i. Selection to increase variance

- $\gamma_{ij} > 0$.   Correlational selection on traits i & j to increase their correlation.

- $\gamma_{ij} < 0$.   Correlational selection on traits i & j to decrease their correlation.

6

Effects of types on selection on changes in G

$$\mathbf{P}^{-1}(\mathbf{P}^* - \mathbf{P})\mathbf{P}^{-1} = \mathbf{P}^{-1}(\mathbf{C} - \mathbf{S}\mathbf{S}^T)\mathbf{P}^{-1}$$

$$= \mathbf{P}^{-1}\mathbf{C}\mathbf{P}^{-1} - (\mathbf{P}^{-1}\mathbf{S})\left(\mathbf{P}^{-1}\mathbf{S}\right)^T$$

$$= \gamma - \beta\beta^T$$

$$\mathbf{G}^* - \mathbf{G} = -\mathbf{G}\beta\beta^T\mathbf{G} + \mathbf{G}\gamma\mathbf{G}$$

$$= -\mathbf{R}\,\mathbf{R}^T + \mathbf{G}\gamma\mathbf{G}$$

$$G_{ij}^* - G_{ij} = -\left(\sum_{k=1}^n \beta_k\, G_{ik}\right)\left(\sum_{k=1}^n \beta_k\, G_{jk}\right) + \sum_{k=1}^n\sum_{\ell=1}^n \gamma_{k\ell}\, G_{ik}\, G_{\ell j}$$

$$= -R_i \cdot R_j + \sum_{k=1}^n\sum_{\ell=1}^n \gamma_{k\ell}\, G_{ik}\, G_{\ell j}$$

7

Change in the genetic variance

$$G_{ii}^* - G_{ii} = -(R_i)^2 + \sum_{k=1}^{n} \sum_{\ell=1}^{n} \gamma_{k\ell} G_{ik} G_{i\ell}$$

Any change in the mean reduces the variance

Note that their can be no direct selection on i ($\beta_i = 0$), but if we have a response, get reduction in variance

Now consider changes in covariances

$$\Delta G_{ij} = -R_i R_j + 2\gamma_{ij} G_{ii} G_{jj}$$

Directional selection in same direction = reduction in covariance

Directional selection in different directions = increase in covariance

Sign of γ_{ij} determines effect of quadratic selection on covariance

Asymmetric responses can occur

Under the standard breeder's equation, the correlated response is $CR_1 = S_2\, h_2\, h_1\, r_A$, so $CR_1 = CR_2$, so long as $S_1 = S_2$

This is no longer true under the infinitesimal model, as the covariances change and which trait (1 or 2) is selected differentially changes the covariance.

$$\Delta G_{ij} = -R_i R_j = \begin{cases} -\beta_1^2\, \sigma_A^2(1) \cdot \sigma(A_1, A_2) & \text{Direct selection on trait 1} \\ -\beta_2^2\, \sigma_A^2(2) \cdot \sigma(A_1, A_2) & \text{Direct selection on trait 2} \end{cases}$$

When disequilibrium-driven selection asymmetries occur, the correlated response will be smaller when selecting on the trait with the higher heritability, as this produces the largest reduction of the genetic covariance.

Changes in G under Gaussian fitness functions

A very common fitness function is the Gaussian

$$W(\mathbf{z}) = \exp\left(-\frac{1}{2}\mathbf{z}^T\mathbf{W}\mathbf{z}\right)$$

If phenotypes and BVs are multivariate normal before selection, they remain so afterwards

$$\mathbf{P}^* = \left(\mathbf{P}^{-1} + \mathbf{W}\right)^{-1}$$

The Bulmer equation has a simple form when using this fitness function,

$$\mathbf{D}_t = \frac{1}{2}\left(\mathbf{D}_{t-1} - \mathbf{G}_{t-1}\left(\mathbf{W}^{-1} + \mathbf{P}_{t-1}\right)^{-1}\mathbf{G}_{t-1}\right)$$

Allele frequency change

Two sources generating genetic covariances

Linkage disequilibrium: alleles with effects on only single traits can become associated, creating a covariance

Pleiotropy: an allele influences two (or more) traits

Changes in G from linkage are transient, as the association decays away once selection stops

Changes in G from pleiotropy are permanent, as the allele frequencies do not change once selection stops

Changes in covariance when allele frequency change occurs

Recall that the heritability provides very little information as to how the variance changes when allele frequencies change

Genetic covariances are <u>even more fragile</u> than genetic variances

Smaller allele frequency change can have larger effects than changes on variance

Hidden pleiotropy: Lots of pleiotropic alleles present but there effects cancel, so no *net* covariance.

complementary pleiotropy: ++ or --

antagonistic pleiotropy: +- or -+

Hidden Pleiotropy: roughly equal numbers of each, no NET genetic correlation.

In this case, allele frequency change can generate either positive or negative genetic correlations.

Thus, a pair of traits with no initial genetic correlation could, in the extreme, either consist of all no pleiotropic alleles or all pleiotropic alleles. Very different outcomes result from these two states.

Models of covariance change under allele frequency change

Hazel, Lush, and Lerner proposed that selection to move two traits in the same direction increases the relative frequency of antagonistic pleiotropic alleles, making genetic covariances more negative as selection proceeds.

++ and -- alleles are quickly lost (or fixed) by selection, while -+ and +- alleles experience less selective pressure, and therefore experience slower allele frequency change

Some experimental support for this:

Friars et al. 1962: Genetic correlations measured from 1949 to 1957 in a series of chicken lines for   production traits.   16 of the 18 correlations showed a negative trend

Genetic Covariances are More Fragile Than Genetic Variances

Classic paper of Bohren, Hill, Robertson (1966)

	Class A	Class B	Class C	Class D
Trait 1	α	β_1	γ_1	0
Trait 2	0	β_2	$-\gamma_2$	δ

$$\sigma(g_1, g_2) = 2\, p_B(1 - p_B)\beta_1\beta_2 - 2p_C(1 - p_C)\gamma_1\gamma_2$$

Theoretical results and simulations showed covariances change more quickly, and more erratically, than variances.

Confirmed the general suggestion by Hazel et al. than eventually genetic covariances generally become more negative. However, depending on the distribution of allele frequencies and effects, the genetic covariance may actually increase in the first few generations.

15

# Can Antagonistic Pleiotropy Maintain Variation?

Several workers suggested that alleles having antagonistic pleiotropic effects on different life-history fitness components (such as reducing fecundity while increasing life span) might be maintained in the population

Same logic, but different outcome from Hazel et al suggestion, which was that such alleles would have longer persistance times, but Hazel et al. did not examine if they would be maintained by selection

Curtsinger et al. Model for this:

	A_1A_1	A_1A_2	A_2A_2
Fitness component 1	1	$1 - h_1\nu$	$1 - \nu$
Fitness component 2	$1 - f$	$1 - h_2 f$	1
Total Fitness	$1 - f$	$(1 - h_1\nu)(1 - h_2 f)$	$1 - \nu$

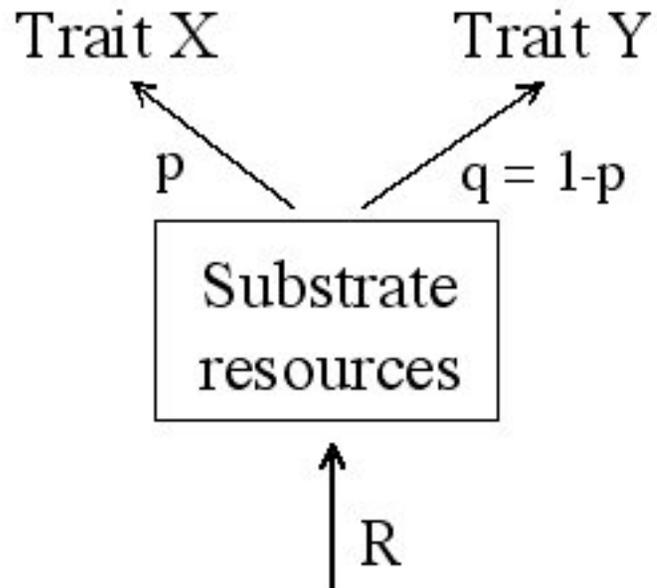$$(1 - h_1\nu)(1 - h_2 f) > \max(1 - \nu, 1 - f)$$
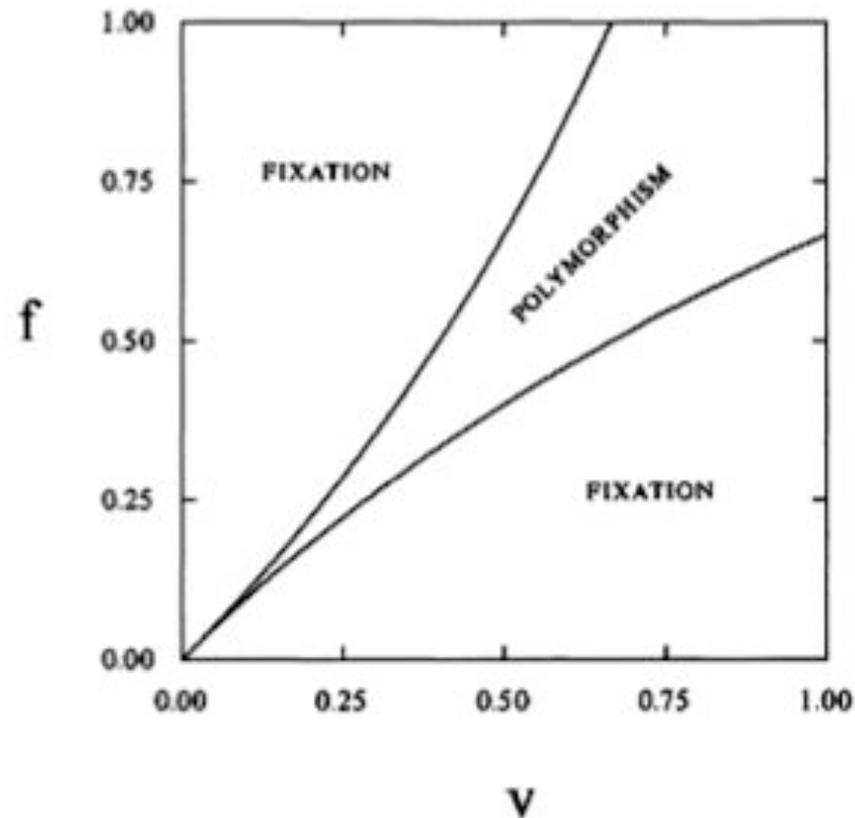
Range of parameters allowing for polymorphism is quite narrow, esp. for weak selection.

For technical reasons, his model likely <u>overestimates</u> the chance of maintaining polymorphisms.

Nature of Pleiotropic covariances

What makes a pleiotropic effect antagonistic vs. complementary?

Resource-partitioning models offer some insight

Alleles contributing to the acquisition of the common resource R are complementary

Alleles contributing to the allocation of this common resource are antagonistic

Obviously, the relative frequencies of these classes of alleles (and their variances) determine if a covariance is positive or negative

18

A simple model for how genetic variation in allocation vs acquisition can generate both positive and negative Correlations. Consider two life history traits
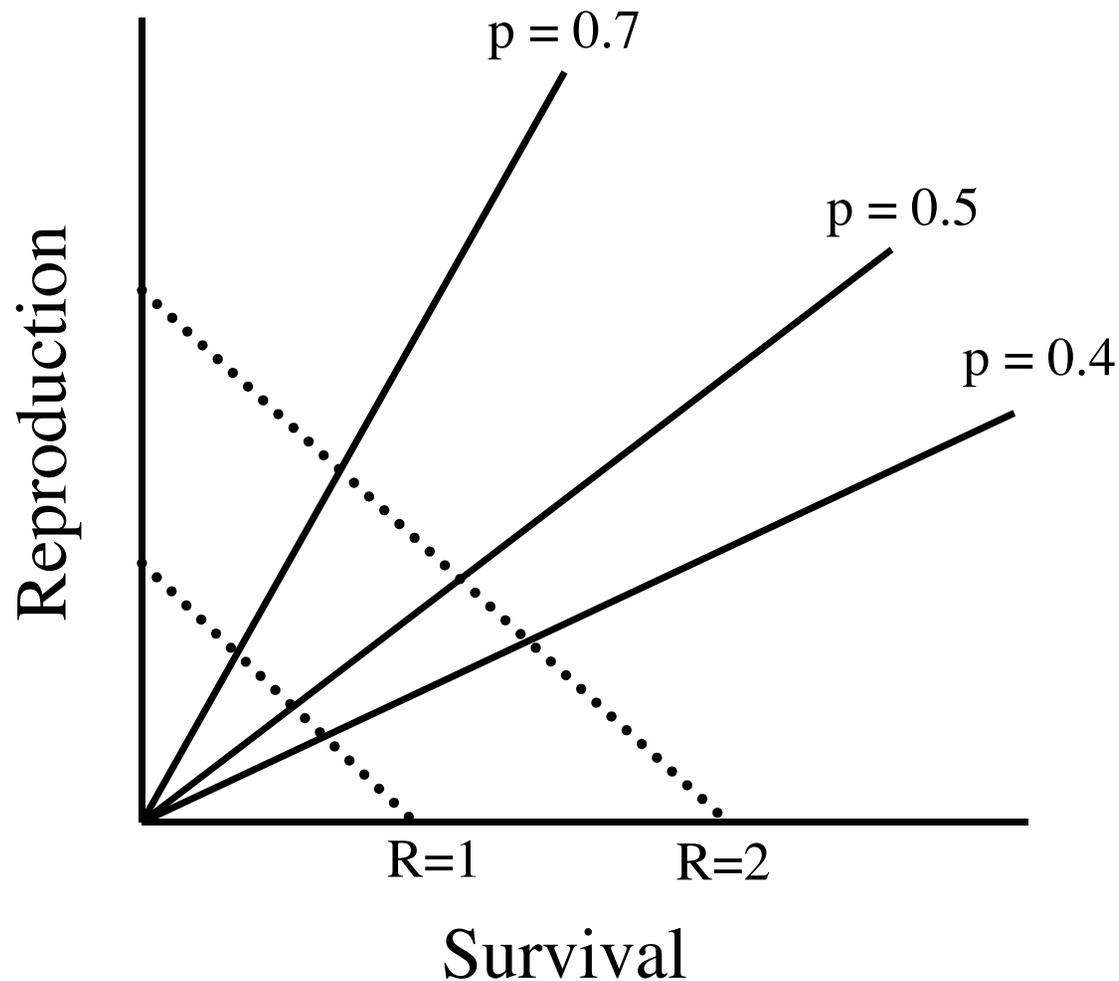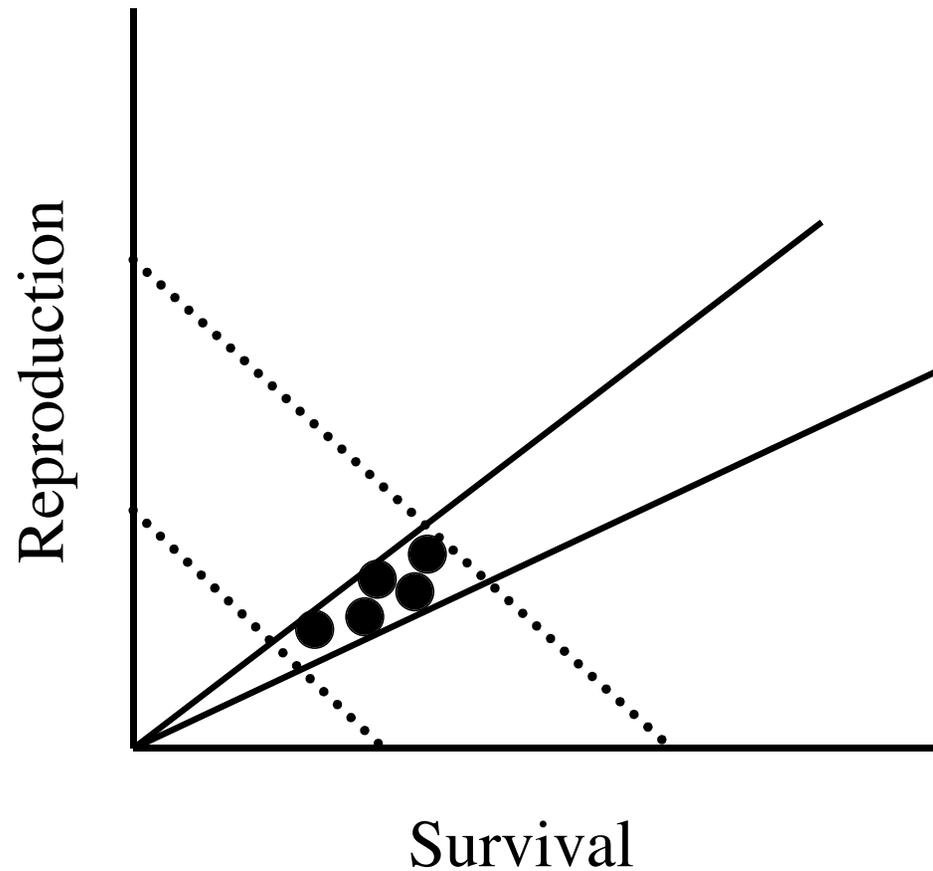
If there is considerable genetic variation in acquisition but much less genetic variation in allocation, this can generate positive genetic correlations between the two traits (here the life history traits of reproduction and survival).
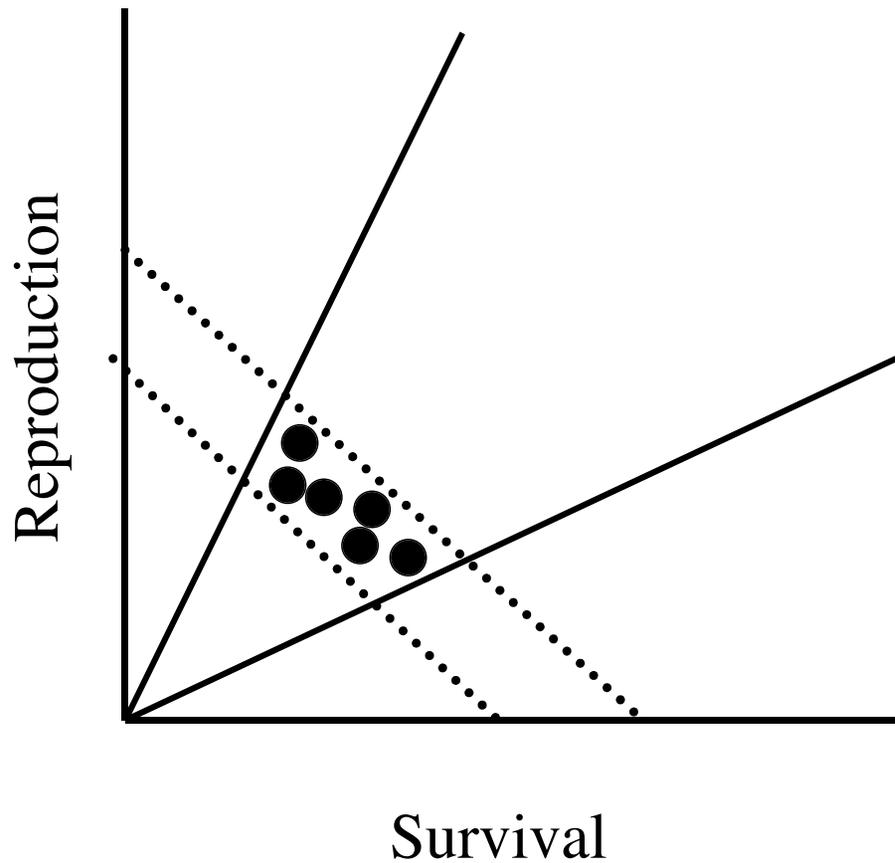
Here small genetic variation in acquisition, larger variation in allocation, generating a negative genetic correlations between the two traits.

Key: Tradeoffs (allocation of the common resource) do no automatically mean negative genetic covariances!

Long term directional selection response: Drift

The simplest model to accommodate allele frequency
Change is the infinitesimal with drift. When
all genetic variance is additive, the expected
value of **G** has a simple form:

$$\mathbf{G}_t = \left(1 - \frac{1}{2N_e}\right)^t \mathbf{G}_0 \simeq e^{-t/2N_e}\,\mathbf{G}_0$$

Hence, the expected cumulative response after t
generations is just

$$\mathbf{R}_t^c = \mu_t - \mu_0 = \sum_{k=1}^{t} \mathbf{G}_k \beta_k = \sum_{k=1}^{t} \left(1 - \frac{1}{2N_e}\right)^k \mathbf{G}_0 \beta_k$$

22

If we assume a constant amount of directional
selection each generation, the cumulative
response simplifies to

$$\mathbf{R}_t^c = \sum_{k=1}^{t} \left(1 - \frac{1}{2N_e}\right)^k \mathbf{G}_0 \boldsymbol{\beta}$$

$$= \mathbf{G}_0 \boldsymbol{\beta} \sum_{k=1}^{t} \left(1 - \frac{1}{2N_e}\right)^k$$

$$\simeq 2N_e(1 - e^{t/2N_e})\mathbf{R}_1$$

The expected total response from the genetic
variation initially present thus becomes

$$\lim_{t \to \infty} \mathbf{R}_t^c = 2N_e\mathbf{R}_1$$

This is the multivariate form of Robertson's
limit: total response is $2N_e$ * initial response vector

Tradeoff: as stronger selection gives initially larger
response but smaller N_e, and hence smaller total response

23

Long term directional selection response: drift and new mutation

Staying within the infinitesimal framework, consider the joint effects of drift removing variation, mutation adding it.

If M denotes the matrix of mutational input, then the mutation-drift equilibrium is

$$\tilde{\mathbf{G}} = 2N_e\mathbf{M}$$

The G matrix at time t thus becomes

$$\mathbf{G}_t = \tilde{\mathbf{G}} + \left(\mathbf{G}_0 - \tilde{\mathbf{G}}\right)e^{-t/2N_e}$$

Putting these together, the cumulative response by generation t becomes

$$\mathbf{R}_t^c = \sum_{i=1}^{t} \mathbf{G}_t \boldsymbol{\beta} = \left[t\,\tilde{\mathbf{G}} + 2N_e(1 - e^{-t/2N_e}) \left(\mathbf{G}_0 - \tilde{\mathbf{G}} \right) \right] \boldsymbol{\beta}$$

We can decompose this two ways. First, the asymptotic response plus the response from the residual component due to the initial variation

$$\mathbf{R}_t^c = t\,\tilde{\mathbf{G}}\boldsymbol{\beta} + 2N_e(1 - e^{-t/2N_e}) \left(\mathbf{G}_0 - \tilde{\mathbf{G}} \right) \boldsymbol{\beta}$$

25

Alternatively, we can decompose the total response into the response from the original variation and the response from new mutation,

$$\mathbf{R}_t^c = 2N_e(1 - e^{t/2N_e})\mathbf{R}_1 + \left(t - 2N_e(1 - e^{-t/2N_e})\right)\tilde{\mathbf{G}}\beta$$

For any particular trait, the ratio of component 2 to the total response is the fraction of response due to new mutation (variation not present at the start of selection).

Long term response: Balance between directional and
stabilizing selection (infinitesimal model results)

One class of models for a selection limit is the
balance between directional (which could be artificial)
and stabilizing selection.

Zeng looked at this problem using a generalized
Gaussian fitness function

$$W(\mathbf{z}) = \exp\left(\mathbf{a}^T\mathbf{z} - \frac{1}{2}(\mathbf{z} - \boldsymbol{\theta})^T\mathbf{W}(\mathbf{z} - \boldsymbol{\theta})\right) = \exp\left(-\frac{1}{2}(\mathbf{z} - \boldsymbol{\psi})^T\mathbf{W}(\mathbf{z} - \boldsymbol{\psi})\right)$$

The matrix **W** describes the nature of quadratic
(stabilizing) selection, while the difference between
the population mean vector μ and the optimal vector ψ describes
the directional selection component

Under this model, the within-generation change in the vector of means and the phenotypic covariance matrix are

$$P^* = \left(P^{-1} + W\right)^{-1} \qquad \mu^* = P^* \left(P^{-1}\mu + W\psi\right)$$

Previous results describe the change in G, while the change in mean is

$$\mu_{t+1} - \mu_t = G_t P_t^{-1} \left(\mu_t^* - \mu_t\right)$$
$$= G_t \left(W^{-1} + P_t\right)^{-1} \left(\psi - \mu_t\right)$$

$$\widehat{\mu} = \psi = \theta + W^{-1}a$$

Zeng's key observation was that when the change in mean equals zero, the solution is independent of the genetic covariance structure **G**.

Long term response: Critically depend on the distribution of allelic effects

The infinitesimal-based models avoid the need to consider the messy genetic details. In reality, they are important.

Model for artificial selection: The traits of interest were in a mutation-(natural) selection balance, and this the variation that forms the foundation for the initial response.

Two general classes of approximations used in mutation-selection models: Gaussian genetic models and House-of-cards (HOC) models.

Gaussian genetic model

Analysis is based on the <u>random-walk mutational model</u>, where the effect a* of a new mutation is given by a* = a + ζ, namely the current value plus a random effect ζ. Results in the effects of selection relative to mutation are weak, many alleles at a locus, distribution of effects roughly normal.

House-of-Cards model

Analysis is based on Kingman's <u>house-or-cards mutational model</u>, where a* = ζ, namely current affect size has no impact on the value in a new mutation. The results are that the effects of selection relative to mutation are strong, few alleles at a locus, most rare. Highly Leptokurtic distribution of allele frequencies, with rare alleles having significant effects

Which model is assumed makes a major difference for long-term response, as simulations by Reeve showed:

Reeve (2000) used simulation studies to examine the consequences of gaussian versus house of cards models for long-term directional response. In his simulations, a population initially under a stabilizing selection-mutation balance has its optimal value shifted by a significant amount (many standard deviations) and the response to selection followed, using the initial estimate of $\mathbf{G}$ (the mutation-selection balance value) to predict response. Reeve found that if allelic effects follow the gaussian model, then the initial $\mathbf{G}$ does a good job of predicting response, despite modest allele frequency changes. If, however, the house of cards model is more appropriate, then the initial $\mathbf{G}$ does a very poor job in predicting response, underestimating the rate of direct and correlated response and overestimating the time to reach the new optimum. This occurs because rare alleles increase in frequency, significantly altering $\mathbf{G}$.

Long term response: Balance between directional and stabilizing selection (finite-locus model results)

Baatz and Wagner considered a two-trait model with directional selection on one trait and stabilizing on the other.

Key result: If pleiotropic alleles are present, they can have a significant impact on the dynamics of the change in mean, even if there is no genetic covariance between traits.

The change in mean turns out to be a function of $\sigma(g_1, g_2^2)$, which can be nonzero, even when there is no correlation among breeding values, $\sigma(g_1, g_2)$, = 0. This can easily happen with hidden pleiotropy.

Consider a favorable rare allele that increases trait 1 and also have an impact on trait 2 (positively or negatively)

As directional selection tries to drive this allele to a higher frequency, it also increases the variance at trait 2, increasing the strength of selection against in. In the extreme, this later selection an be sufficiently strong as to stop directional selection

Baatz and Wagner call this the <u>Pooh effect</u>, after Winnie the Pooh eating to much honey and getting stuck in rabbit's house

The key message from the above analysis is that while genetic covariances are excellent for providing a quick snap-shot of the local dynamics, when allele frequency change occurs, all bets are off. Thus, one can have a situation where selection on an unobserved trait that is *not* genetically correlation with our focal trait can still influence its fitness if pleiotropy is present and allele frequency change occurs.

Long term response: Stabilizing selection
Mutation-selection balance

Much work has been done on mutation-selection balance problems with stabilizing selection on a trait (or traits) removing selection while mutation introduces it.

Problem: The simple models don't work: Selection in the wild is too strong to account for the high levels of genetic variation seen

Classic analysis: Kimura-Lande model assuming Gaussian distribution of allelic effects at each locus

Under gaussian stabilizing selection, we have

$$\Delta \mathbf{G} = -\mathbf{G}\,\omega\,\mathbf{G}, \quad \text{where} \quad \omega = \mathbf{W}^{-1} + \mathbf{E}$$

The equilibrium G matrix is the sum over all loci,

$$\tilde{\mathbf{G}} = 2\sum_{i=1}^{n} \omega^{1/2} \left(\omega^{-1/2}\mathbf{M}_i\omega^{-1/2}\right)^{1/2} \omega^{1/2}$$

$$= 2\omega^{1/2} \left[\sum_{i=1}^{n} \left(\omega^{-1/2}\mathbf{M}_i\omega^{-1/2}\right)^{1/2}\right] \omega^{1/2}$$

If there is no net pleiotropy, we have

$$\tilde{\mathbf{G}} = 2\omega^{1/2} \left[\sum_{i=1}^{n} (c_i\mathbf{I})^{1/2}\right] = k\omega^{1/2}$$

$$\tilde{\mathbf{G}} = 2\omega^{1/2}\left[\sum_{i=1}^{n}(c_i\mathbf{I})^{1/2}\right] = k\omega^{1/2}$$

A neat result from linear algebra is that the eigenvectors of $\mathbf{A}^{1/2}$ are the same as the eigenvectors of $\mathbf{A}$. Hence, the above result suggests that G wil evolve such that its axis of variation mirror those of g. Further, since the eigenvalues of $\mathbf{A}^{1/2}$ are just the square roots of the eigenvalues of $\mathbf{A}$, the variation along these axis is also proportional.

Are the axes of $\mathbf{G}$ and the quadratic fitness surface similar? i.e., can $\mathbf{G}$ evolve to have axes similar to γ?

Common orientation of γ and G??

Are the axes of **G** and the quadratic fitness surface similar? i.e., can **G** evolve to have axes similar to γ?

Hunt et al. Male call components in the cricket Teleogryllus commodus

	$\mathbf{g}_{max}$	$\mathbf{g}_2$	$\mathbf{g}_3$	$\mathbf{g}_4$	$\mathbf{g}_5$
% Variation	51.0	25.6	12.9	6.9	3.6
$\mathbf{g}_i^T \gamma \mathbf{g}_1$	0.005	-0.012	-0.051	-0.097	-0.011
$\mathbf{g}_i^T \Delta \mathbf{G} \mathbf{g}_1$	0.00400	-0.00283	-0.00285	-0.00212	-0.00009
% change in Genetic SD	6.6	-7.8	-11.0	-13.0	-3.8

The key observation of Hunt et al. is that the amount of genetic variation along a particular direction decreases with the strength of stabilizing selection along that axis (genetic variation decreases, and the strength of stabilizing selection increases, as we move from $\mathbf{g}_2$ to $\mathbf{g}_4$). While the fifth (and final) PC shows a departure from this pattern, given there is no support for this PC, estimates associated with it must be viewed with some caution.

Model Assumptions, Genetic Correlations, and Hidden Pleiotropy

Two types of correlations, phenotypic and genetic, enter into our discussion of multivariate models. Under Gaussian *fitness* models, selection on one trait reduces the phenotypic variance of any phenotypically-correlated traits (Equation 31.13a), which imparts selection on the second trait. Genetic correlations also influence selection response and can be generated by pleiotropic mutations. Under the gaussian genetic assumption, if there are no correlations among pleiotropic mutations and no phenotypic correlations in the selection function (**M** and **W** are diagonal), then the equilibrium amount of variation maintained at each trait is independent of selection at the others (Lande 1980, 1984; Turelli 1985). However, under the HC approximation if there is hidden pleiotropy, then selection on a trait can be significantly influenced by selection on uncorrelated (genetically and phenotypically) traits (Turelli 1985, 1988b; Wagner 1989; Slatkin and Frank 1990). With hidden pleiotropy, the equilibrium genetic variance that a trait can maintain is significantly reduced by selection on uncorrelated traits. If the results from the HOC approximation are a good reflection of the true biological reality, then selection from uncorrelated traits can significantly impact the long-term behavior of traits.