

This manuscript is originally published in “The Open Anthropology Journal” Volume 3, 2010; 206-229pp.

We offer this L^AT_EX version since it is easier to read. If you want the journal-formatted version, please go here: <http://www.bentham.org/open/toanthj/openaccess2.htm>

PLEASE CITE AS:

Lawler RR, Blomquist GE (2010) Multivariate selection theory in primatology: An introduction to the concepts and literature. Open Anthropology Journal. 3: 206-229.

Multivariate selection theory in primatology: An introduction to the concepts and literature

Richard R Lawler

Department of Sociology and Anthropology
James Madison University, USA

Gregory E Blomquist

Department of Anthropology
University of Missouri, USA

October 15, 2010

Contents

1	Introduction	5
2	Multivariate phenotypic selection models	6
3	A Model of Phenotypic Variation	13
4	Fitness, Selection, and Adaptation	15
5	Model assumptions and the role of models in science	18
6	Extensions to specific evolutionary phenomena	21
6.1	Sexual selection and mate choice	21
6.2	Sexual selection and size dimorphism	23

6.3	Life history theory	25
6.4	Evolutionary constraints	28
6.5	Allometry	29
6.6	Ecological morphology	31
6.7	Social behavior	32
7	Quantitative Genetic Studies of Primates: Some Examples	35
7.1	Measuring selection in primate populations	35
7.2	Measuring heritability and genetic correlations in primate populations	37
8	Conclusions	42
9	Acknowledgements	43

Abstract

With its focus on phenotypes, inheritance, and natural selection, evolutionary quantitative genetics constitutes a bridge between the genetic architecture of traits and evolutionary dynamics. This field has produced a vast theoretical literature with numerous empirical studies supporting the basic theoretical principles of quantitative genetics. At minimum, quantitative genetic studies require data concerning phenotype, fitness, and kin relations for individuals in a population. These data are rarely available to primatologists who study primate populations in the wild. Increasingly, however, evolutionary quantitative genetic studies have been conducted on wild and free-ranging primates, and several long-term studies of wild primates are producing the necessary data to conduct quantitative genetic studies. Our goal in this review is to provide a thorough and (hopefully) gentle introduction to quantitative genetic theory, with particular emphasis on multivariate selection theory. We review the basic steps in deriving a multivariate equation for evolutionary change and we then show how this basic equation can be modified in order to study sexual selection, life history theory, evolutionary constraints, allometry, ecological morphology and social behavior. We also discuss the epistemological role of quantitative genetic models as well as basic concepts such as fitness, selection, and adaptation as they pertain to quantitative genetic studies. Finally, we review some recent quantitative genetic studies of wild and free-ranging primates.

1 Introduction

Adaptation, no doubt, is a cornerstone of biological anthropology in general and of primatology in particular. A lot of research seeks to generate or test hypotheses about the adaptive basis of a particular trait or behavior in primate populations. Whether one is studying agonism among adult male rhesus macaques or the origins of bipedalism, such topics are almost always studied with reference to the fitness consequences of a particular behavioral strategy or anatomical configuration. For the most part, insights into the adaptive evolution of primate behaviors and anatomical traits come from three approaches: 1) long-term ecological investigations of one or a few primate social groups in the wild; 2) measuring osteological specimens in a museum in order to understand functional aspects of anatomical form; or, 3) developing interspecific comparisons of primate traits and a putative selective pressure—the comparative method. A bit of reflection about how these three approaches are carried out, suggests that they operate on different levels of the biological hierarchy and draw from different sample sizes. Primatologists in the field often follow focal groups or focal animals resulting in numerous datapoints collected on about 5-25 animals. Anatomists in the museum, on the other hand, often take homologous measurements on anatomical landmarks from various species resulting in a sample size that is often dependent on the number of specimens available in the museum collections. Finally, comparative studies often use an average value of a particular trait and map this trait against a putative selective pressure across a phylogeny; here, the sample size depends on which species are included in the analysis. It is interesting that these approaches all focus on adaptation, but do not directly study the unit of evolution in which adaptations are forged—the population. Given that the population is the unit of evolution it would be illuminating to gather data on variation among individuals in fitness and phenotype as well as some information on the fidelity through which traits and behaviors are transmitted across generations. If one could collect such information, then one could gain insight into the processes by which natural selection crafts adaptations across generations. That is, one could estimate the strength of selection acting on traits or behaviors as well as the heritability of these traits or behaviors. Such an endeavor is easier said than done. Primates, as an order, tend to live longer and reproduce later in life than other orders of mammals. Such a delayed life history schedule makes collecting data on variation in phenotypes and fitness rather difficult. In order to obtain sufficient data to estimate things like selection coefficients and trait heritabilities long-term data are required from animals with known fates and kinship.

In this paper, we outline widely used techniques that allow one to quantify the action of natural and sexual selection and to estimate the heritable portion of the phenotype. These techniques have been widely applied in experimental and/or wild populations of insects, fish, birds, and some mammal species. Only rarely, however, are they applied to wild or free-ranging primate populations. The data required to achieve these goals entails considerable effort to collect, but at present, there are numerous long-term studies that have collected data on genetic relationships, reproduction, and phenotypic variation. Many of these long-term studies have recently been discussed by Strier et al., [90] [91] in the context of a comparative primate life history database. Thus with the continuing accrual of data from wild and free-

ranging primate populations and new analyses from long-term captive populations, we feel that the techniques and methods we outline below will become more prevalent in primatology. In short, numerous studies are beginning to incorporate genetics and live-capturing into their research program; the methods outlined below will hopefully guide these studies in the analysis of these new datasets.

Our goal is to outline the body of theory for study of adaptive phenotypic evolution from a quantitative genetic perspective. By “quantitative genetic”, we mean that the traits under study are continuous and quantitatively coded by many loci. The models we discuss are sometimes called “phenotypic selection models,” since they focus on how phenotypes change under selection in a multi-trait (i.e., multivariate) context. Usually, the only genetic information one needs in such models is information on the faithfulness through which traits are transmitted across generations. As such, the power of these methods is that they provide insight into the processes of adaptive evolution while accounting for the fact that traits within an organism are sometimes integrated and must be analyzed as a unit. In this paper we review the basic components of multivariate selection theory. Though we rarely discuss primates per se, our goal is to provide an accessible introduction to primatologists, with the hope that these techniques will be adopted by them. Those who already possess some knowledge of these techniques will find our review excessively wordy. In short, we would rather insult the intelligence of the expert than alienate the non-initiate.

2 Multivariate phenotypic selection models

To begin, we will consider selection on a single trait. In what follows, we use the term “trait” to refer to some aspect of the phenotype (e.g., height, probability of living to age x , body mass) and “trait value” to the corresponding value that the phenotypic trait can take in an individual (e.g., 166mm, 0.95, 68kg). Table 1 gives the definitions of the terms used in the equations. Consider a trait such as height. Let’s denote height as trait z_1 . Selection acts on individuals and as a consequence, the population evolves; in this case, we want to track phenotypic change at the population-level. Specifically, we want to understand how the mean of z_1 changes from generation to generation due to selection. We will denote the mean change in trait z_1 as $\Delta\bar{z}_1$. The information we would need to understand this is the following: which trait values are being favored by selection (Is it tall people...short people, etc?) and how much of a parent’s height is passed on to offspring?—or put another way, how heritable is height? Imagine a hypothetical population who’s average height is 5’8”. Then, imagine we had knowledge that all people of height 6 feet or greater in this population were allowed to randomly breed with each other. Further, say we had some information that height was heritable. Now, armed with this knowledge, imagine we examined the distribution of height of the offspring produced by the tall parents (i.e., those selected to breed) and we found it to be, on average 5’11”. In thinking about this information, we could conclude something like,

Table 1: Description of mathematical terms used in this paper

Variables and terms	Description
z	Any phenotypic trait (subscripts will be used to distinguish different traits)
\mathbf{z}	A vector of traits z , e.g., z_1, z_2, \dots, z_n
$\Delta \bar{z}$	The average change in trait z between generations
$\Delta \bar{\mathbf{z}}$	The average change in vector of traits z_n between generations
S	The selection differential
V_A	Additive genetic variance
V_P	Phenotypic variance
α	Breeding value
h^2	Narrow sense heritability
$V_P(1, 1)$	The phenotypic variance in trait 1
$C_P(1, 2)$	The phenotypic covariance between traits 1 and 2
$V_A(1, 1)$	The additive genetic variance in trait 1
$C_A(1, 2)$	The additive genetic covariance between traits 1 and 2
$V_E(1, 1)$	The environmental variance in trait 1
$C_E(1, 2)$	The environmental covariance between traits 1 and 2
ρ_P	The phenotypic correlation between traits
ρ_G	The additive genetic correlation between traits
β	The directional selection gradient
$\gamma_{i,i}$	The nonlinear selection gradient for trait i
$\gamma_{i,j}$	The correlational selection gradient among traits i, j
\mathbf{P}	The phenotypic (co)variance matrix
\mathbf{G}	The genetic (co)variance matrix
\mathbf{A}	The projection matrix
$\boldsymbol{\beta}$	Vector of selection gradients
\mathbf{s}	Vector of selection differentials
λ	A measure of fitness measured as the rate of population growth
R_o	A measure of fitness measured as per-generation rate of offspring production
$\boldsymbol{\theta}$	A vector of optimum trait values
$\boldsymbol{\omega}$	A matrix of stabilizing selection coefficients
\mathbf{B}	Additive genetic covariance matrix between males and females
$\nabla \lambda$	Vector of directional selection coefficients for life history traits
$\frac{\partial \lambda}{\partial z_k}$	Sensitivity, or directional selection coefficient for a life history trait
$\boldsymbol{\Psi}$	Matrix of interaction coefficients in social selection models
$\mathbf{C}_{\alpha\mathbf{p}}$	Matrix of correlation coefficients between breeding value and phenotype

The change in average height due to selection across generations is equal to the phenotypic advantage of those people being selected for multiplied by the degree to which height is heritable.

Thus if tall people were being selected for (i.e., allowed to out-reproduce short people) and, on average, tall parents had tall children (i.e., height is heritable), then we would know the change in average height across generations is due to selection for tall people. To represent our word equation mathematically, we could write

$$\Delta \bar{z}_1 = Sh^2 \tag{1}$$

S is called the *selection differential* and represents the difference in average height between the tall people and the total population. Say, the average height of the tall people was 6'2". Thus S is simply

$$S = 6'2'' - 5'8''$$

The term h^2 in equation 1 refers to the heritability. Formally, it is the ratio of the additive genetic variance (V_A) to the phenotypic variance (V_P) ($h^2 = V_A/V_P$). Heritability is a number between 0 and 1; when heritability is equal to 1, then all of the variation in trait values for a particular trait is due to additive genetic variation. Heritability determines the degree to which offspring resemble their parents based on additive genetic effects, because it includes the term V_A . Additive genetic variance is the genetic variation in the population that produces the genetic-based similarity between parent and offspring phenotypes (we will expand on this concept in Section 3). If we assume the heritability of height is 0.5, then we can calculate the change in average height using equation 1. We know that $S = 6''$, thus multiplying this number by the heritability yields 3 inches. In this regard, the next generation of offspring will be three inches taller, on average, due to selection for tall parents ($5'8'' + 3'' = 5'11''$).

Substituting V_A/V_P for h^2 and rearranging, we can rewrite equation 1 as

$$\Delta \bar{z}_1 = S \frac{V_A}{V_P} \quad (\text{equivalently}) \quad \Delta \bar{z}_1 = V_A \frac{1}{V_P} S \tag{2}$$

The reason we rewrote equation 1 as equation 2 was that this equation will have some correspondence with equations we develop below. In equation 2, the mean change in a trait value is proportional to the selection differential (S) and additive genetic variance (V_A), and inversely proportional to the phenotypic variation. Equation 2 makes it clear that V_A is what matters when calculating the response to selection. If V_A were zero, then the entire right side of equation 2 would be zero and there would be no response to selection. Thus we can conclude that in order to see some change in average phenotype across generations, we need to have some amount of genetic variation that produces a correspondence between offspring and parents. Also note that if we scale S by phenotypic standard deviation ($\sqrt{V_P}$), we get what is known as the *selection intensity*, usually noted as i ($i = S/\sqrt{V_P}$).

The selection differential is defined as the difference in average trait values between those allowed to breed and the total population. Implicitly, what we were doing was stating that

the selection differential involves a non-zero covariance between fitness and phenotype. That is, we were stating that all folks whose height is equal to or greater than 6' were allowed to breed (they have a high fitness) whereas those shorter folks have lower fitness because they are not allowed to breed. The selection differential in this case represents *truncation selection* since there is some cut-off point that determines who breeds and who doesn't. Obviously, in nature, selection rarely works with such fastidiousness. More often, we would likely see a distribution of offspring production that is associated with a distribution of trait values. In such cases, the selection differential is calculated in a slightly different way than the case for truncation selection; it is no longer the difference in average trait values among those selected and the total population; rather, S is simply the relationship, itself, between relative fitness and phenotype (prior to reproduction). That is, in nature, S is measured as the covariance between fitness (w) and phenotype. More formally then,

$$S = (\text{Cov}(w, z_i)) \quad (3)$$

Any association between a particular trait and fitness can be measured using equation 3. Here, fitness is scaled by mean fitness ($w = W/\bar{W}$). The selection differential, as defined in equation 3, measures the *total* selection on a trait even if the trait is phenotypically correlated with another trait.

What if selection acts on multiple traits? How do the traits change across generations in such cases? To model this, we need to introduce some new traits and terms. In addition to trait z_1 , let's also consider trait z_2 . We can model selection acting on these two traits in the same way that we modeled it for trait z_1 . However, the new thing that we need to account for is the fact that the traits may be correlated (i.e., they covary). For example, if trait z_2 is "width" (say, a measurement across the hips), we need to account for the fact that height (z_1) and width (z_2) phenotypically covary. (We denote the *phenotypic* covariation between traits 1 and 2 as $C_P(1,2)$; following from this, we note that any trait which covaries with itself is simply the variance in that trait, for example, the variance in trait 1 is denoted $V_P(1,1)$). Imagine that the two traits are indeed positively correlated, but that selection is acting only on height (how we would know is this discussed below, p. 10). In this case, we expect that selection is acting to make people taller, but that as a consequence of the positive phenotypic covariation, there is also selection for increased width. If we calculated the selection differential, S , for height (i.e., S_1) we would find that it is positive (since tall people have high fitness) but similarly, if we calculated S for width (i.e., S_2), we would also find it to be positive. However, the reason S_2 is positive it due to the fact that width phenotypically covaries with height, which is under selection. Put another way, selection for increased height is due to direct selection acting on height (taking into account the amount of phenotypic variation in height) plus the indirect selection acting on width (taking into account the phenotypic covariation between height and width). Recall the S measures total selection (direct selection plus indirect selection due to phenotypic covariation) acting on a trait. Thus we can write the selection differential for height (S_1) as

$$S_1 = \text{direct selection on } z_1 \times V_P(1,1) + \text{indirect selection on } z_2 \times C_P(1,2) \quad (4)$$

In a similar fashion we can write out S_2 as

$$S_2 = \text{direct selection on } z_2 \times V_P(2,2) + \text{indirect selection on } z_1 \times C_P(2,1) \quad (5)$$

Now let's introduce some new notation. Let's make β_i denote direct selection on trait i ; the β coefficients are known as *selection gradients*. Thus we can write equations 4 and 5 as,

$$S_1 = \beta_1 V_P(1, 1) + \beta_2 C_P(1, 2) \quad (6)$$

and

$$S_2 = \beta_2 V_P(2, 2) + \beta_1 C_P(2, 1) \quad (7)$$

Now, note that $\beta = \frac{1}{V_P} S$. That is, direct selection on a trait is equal to the reciprocal (or inverse) of phenotypic variation multiplied by the selection differential. When we extend this notion to two or more traits, we get a conceptually similar, but not mathematically identical, relationship to $\beta = \frac{1}{V_P} S$. In the two trait case, equations 6 and 7 constitute a system of equations and thus solving for the β terms must be done simultaneously. Intuitively, however, we know that when dealing with multiple traits it is necessary to account for phenotypic covariances among traits. Thus we would expect that direct selection acting on multiple traits involves not only selection differentials acting on each trait but also takes into account correlations among traits. For traits 1 and 2, we have

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} V_P(1, 1) & C_P(1, 2) \\ C_P(2, 1) & V_P(2, 2) \end{pmatrix}^{-1} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \quad (8)$$

In equation 8 we have expressed the direct selection coefficients (the β coefficients) and the selection differentials (S) as vectors (columns of numbers) and associations among traits is captured by a matrix (a box of numbers) that contains information on the phenotypic associations. Formally, the “ -1 ” denotes the inverse of this matrix. Hence the matrix in equation 8 is the inverse of the phenotypic variance-covariance matrix. In more compact notation, equation 8 can be represented as

$$\boldsymbol{\beta} = \mathbf{P}^{-1} \mathbf{s} \quad (9)$$

where boldface terms denote vectors or matrices. The \mathbf{P}^{-1} term represents the inverse of the phenotypic covariance matrix, but this is *not* the same thing as a matrix of the reciprocals of variance/covariance of each trait. Written out, equation 9 is the following (which emphasizes that the entries in the \mathbf{P}^{-1} matrix $\neq 1/V_P$)

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \frac{1}{V_P(1, 1)V_P(2, 2) - C_P(1, 2)C_P(2, 1)} \begin{pmatrix} V_P(2, 2) & -C_P(1, 2) \\ -C_P(2, 1) & V_P(1, 1) \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \quad (10)$$

If we had three traits under study then equation 8 would be expanded as such

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} V_P(1, 1) & C_P(1, 2) & C_P(1, 3) \\ C_P(2, 1) & V_P(2, 2) & C_P(2, 3) \\ C_P(3, 1) & C_P(3, 2) & V_P(3, 3) \end{pmatrix}^{-1} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad (11)$$

Equation 11 represents the selection differentials, selection gradients and phenotypic variances/covariances in terms of vectors and matrices for three traits. Note that we can write out the direct action of selection on trait 1 for the three trait case from equation 11 as follows

$$\beta_1 = V_{P11}S_1 + C_{P12}S_2 + C_{P13}S_3 \quad (12)$$

In this case, terms such as V_{P11} and C_{P12} represent the row by column positions within the \mathbf{P}^{-1} matrix (e.g., V_{P11} is in the first row and first column of the \mathbf{P}^{-1} matrix and this value is multiplied by S_1). And as before, equation 11 can also be written in a much more convenient format; thus

$$\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{s} \quad (13)$$

Equation 13 (or 11) above provides a description of how to parse selection into components of direct and total selection. The right side of equation 13 is similar—but not homologous—to the $\frac{1}{\bar{v}_P}S$ term in equation 2 (since the inverse of a matrix is not a simple matter of taking the reciprocal of each entry). The β coefficients, or selection gradients, are a key component in the theory of multivariate phenotypic selection models—they measure the direct selection acting on traits while controlling for the effects of indirect selection. They help us identify the true target of selection. In our example above concerning height and width, both of the selection differentials, S_1 and S_2 , would provide evidence for selection on these traits. However, only by measuring the selection gradients would we know that selection is targeting height and that the selection on width is indirect (i.e., $\beta_1 > 0, \beta_2 \approx 0$). Armed with the knowledge that selection is targeting height, not width, we could go out and look for the ecological or social reason as to why there are fitness differences with respect to height. In short, the β coefficients provide a powerful way to quantify the direct action of selection on traits when studying multiple traits.

So far, we have discussed ways to disentangle how selection acts on multiple traits that phenotypically covary. However, what we would like to know is how multiple traits respond to selection across generations. Thus we seek an analogous term for V_A in equation 2. Since we are considering multiple traits, we want to focus on trait covariances as well as variances. We do this using an additive genetic variance-covariance matrix for the traits under consideration. For traits z_1, z_2 , and z_3 , we can represent their additive genetic variation ($V_A(i, i)$) and covariation ($C_A(i, j)$) as

$$\mathbf{G} = \begin{pmatrix} V_A(1, 1) & C_A(1, 2) & C_A(1, 3) \\ C_A(2, 1) & V_A(2, 2) & C_A(2, 3) \\ C_A(3, 1) & C_A(3, 2) & V_A(3, 3) \end{pmatrix} \quad (14)$$

This matrix is known as the \mathbf{G} matrix. It is another key component of multivariate selection theory. The \mathbf{G} matrix tells us the degree to which traits in offspring will resemble the traits in their parents, while also noting that some traits may be genetically correlated. That is, traits that have additive genetic covariation will respond jointly to selection acting on one trait. Common mechanistic explanations for genetic covariation among traits is due to pleiotropy (where one genetic locus controls two or more different traits) or linkage disequilibrium (where two different alleles controlling two different traits are co-inherited

more often than expected by chance, thus the two traits tend to respond to selection as a unit rather than independently). In general, pleiotropy is thought to be more ubiquitous than linkage in maintaining genetic correlations, since recombination will generally break down correlations due to linkage disequilibrium. However, linkage disequilibrium is important in many models of sexual selection; further, genetic drift, non-random mating, as well as selection can maintain linkage disequilibrium in populations. Informative discussions of the evolutionary implications of linkage disequilibrium are found in Gillespie [46] and Rice [80]. One important note about the \mathbf{G} matrix is that it assumes that all traits potentially covary but it doesn't account for the fact that some covariances may be more important than others, or that some covariances may not exist at all. Path analysis is a technique that allows one to test for hierarchical or complex covariance structure among traits. To develop path analysis models, one can draw from ecological, physiological, or biomechanical principles to specify the covariance among traits. Rice [80] gives a short but clear description of path analysis for phenotypic selection models.

The \mathbf{G} matrix is analogous to the V_A term in equation 2, and since we know that β is analogous to the $S\frac{1}{V_P}$ term, we can write the response to selection acting on three (i.e., multiple) traits as

$$\begin{pmatrix} \Delta\bar{z}_1 \\ \Delta\bar{z}_2 \\ \Delta\bar{z}_3 \end{pmatrix} = \begin{pmatrix} V_A(1,1) & C_A(1,2) & C_A(1,3) \\ C_A(2,1) & V_A(2,2) & C_A(2,3) \\ C_A(3,1) & C_A(3,2) & V_A(3,3) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \quad (15)$$

As before, we can use matrix notation to represent equation 15 more compactly. Thus,

$$\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta} \quad (16)$$

and noting equation 13 above, this equation can also be written out as

$$\Delta\bar{\mathbf{z}} = \mathbf{G}\mathbf{P}^{-1}\mathbf{s} \quad (17)$$

Equation 16 (or 15 or 17) is one of most famous and notorious equations in evolutionary genetics. It was derived by Lande in 1979, although previous researchers had also derived it in somewhat different form [95]. It allows one to study how the mean value of multiple traits respond to selection across generations as a result of direct selection on each trait (section 5 lists the assumption that go into this equation). Analogous equations can be derived for stabilizing and disruptive selection [60]. We don't explicate these latter equations, but in section 4 we describe and graphically depict these different types of selection. An important property of equation 16 is that a trait can still change across generations even though it does not experience direct selection. In the notation of equation 15, it is possible to observe $\Delta\bar{z}_2 > 0$ but not observe $\beta_2 > 0$. This occurs because trait 2 is genetically correlated with some trait that does experience selection, for example: $C_A(1,2) > 0$ and $\beta_1 > 0$. Though our treatment above may seem complicated to the newcomer, it is actually quite superficial. The motivated reader should consult Lande [57], Lande and Arnold [60], Phillips and Arnold [73], Arnold, [6]; Brodie et al., [15], and Walsh [95] for further details and derivations.

3 A Model of Phenotypic Variation

In this section we want to look at a simple model of the phenotype. Through our discussion we will introduce numerous terms that will be useful in the rest of this paper. We begin by discussing an “intuitive” model of the phenotype (P), namely that a phenotype is built from genetic factors (G) and environmental factors (E). If these are the only two factors that make up the phenotype, then we can write, $P = G + E$. We will assume that the trait is continuous (e.g., height) and that many genetic loci each contribute a small amount to height. If we were looking at the trait in a bunch of individuals, we could look at the variation in this phenotype (V_P) and, similarly, express the phenotypic variation as a the sum of the genetic and environmental variation:

$$V_P = V_G + V_E. \tag{18}$$

V_E represents the amount of phenotypic variation that is due to environmental variation (for example, when some of the variation in height in a population is determined by diet, not genes). In equation 18 we are explicitly assuming that there is no covariance between genetic and environmental variation; when such covariances exist equation 18 would be written as $V_P = V_G + V_E + 2C_{GE}$, where the C_{GE} term represents the covariation between genotype and environment. In looking at equation 18, it should be obvious that the terms “genetic” and “environmental” actually gloss over a fair amount of complexities. Depending the question at hand, it is possible to further break down each of these terms into something more specific. For example, genetic factors (G) can influence a phenotype through the factors that cause parents to resemble offspring, which are often known as additive effects, through dominance (where one allele masks the phenotypic effect of another another allele at a locus), and through epistasis (different genetic loci interact in a complex manner to produce a phenotypic effect). Thus the genetic factor (G) can be broken down into additive (A), dominance (D), and epistatic (I) factors. And as before, we are considering a population of phenotypes so we are more interested in the variation among phenotypes (V_P) and the variation in these factors. In this case, we can write variation in genetic factors in terms of variation due to dominance, epistasis, and additive effects (and we assume that there is no covariation among these terms):

$$V_G = V_A + V_D + V_I \tag{19}$$

In a similar fashion, environmental factors can also be redefined to capture more realism. Consider that some of the phenotypic variation in a population is due to the fact that every individual who was born in a particular year incurred the same environmental conditions (e.g., they all experienced drought-like conditions for their first three years of life thereby stunting growth). In this case, some of the phenotypic variation would be due to birth-year effects, or cohort effects (V_C). Similarly, some sets of siblings in a population may all be very healthy because their mother was a very healthy and an attentive parent. Here, some of the variation in phenotype would be due variation in mothering ability and/or the

placental environment provided by maternal effects (V_M). In fact, it is possible to divide up environmental (and genetic) factors into almost any conceivable component of variation that is relevant. Finally, it is important to include a term that is often called “residual variation (V_R);” this term captures any remaining phenotypic variation that cannot be attributed to one of the other terms in our model of phenotypic variation. Taking all the terms together, our model of phenotypic variation is the following:

$$V_P = V_A + V_D + V_I + V_C + V_M + V_R. \quad (20)$$

It should be noted that this partitioning of phenotypic variation is only one possible way to model phenotypic variation. In practice, most researchers only include those terms they are interested in and/or those terms for which their are available data (e.g., to model dominance effects requires pedigree data that includes both paternal and maternal links). In the same sense that variation in phenotype (V_P) can be broken down into genetic and environmental components, so can the covariation between two traits: $C_P(1, 2) = C_A(1, 2) + C_E(1, 2)$ (which again assumes no covariation between genotypes and environment). While additive genetic covariances ($C_A(1, 2)$) are produced by linkage or pleiotropy, the environmental covariance between traits ($C_E(1, 2)$) is essentially a term that means “everything else” that is not heritable and not modeled by a specific covariance term. In this regard, two traits can environmentally covary due to a shared external circumstances (e.g., the traits are similar due to experiencing a common temperature).

In order to understand patterns of inheritance, we need to find the portion of genetic variation that causes offspring to resemble their parents—this is known as additive genetic variation (V_A). We single out V_A from other genetic components because only V_A is responsible for the resemblance between parent and offspring (Futuyma, [45] pages 413-414, gives a clear example of why V_A produces a correspondence between parent and offspring); other types of genetic variation, for example dominance variation (V_D), can influence phenotypic variation but this factor is not necessarily transmitted across generations since genotypes are not inherited from parents in diploid, sexually reproducing species (only alleles are inherited). While we are considering parent-offspring covariances, two additional points are worth mentioning here. First, it is important to remember that non-genetic factors such as maternal effects (and other shared environments) can produce similarity between parents and offspring; second, there is a form of epistatic interaction known as “additive by additive” epistasis, and this form of epistasis is also transmitted across generations (a good explanation of epistasis is given in [80]). Ultimately, a lot of quantitative genetics research is concerned with estimating the various components that make up V_P , with particular attention heaped on V_A . This is because V_A is a key term in the heritability (h^2), which is given in equations 1 and 2 and also comprises the terms in the \mathbf{G} matrix (equation 14). It is also a key factor that determines part of the selection response in equation 16. Note that when we are considering the phenotype of a single individual (e.g., a parent), the additive genetic component of the individual’s phenotype is called the “breeding value” (α). For simplicity, one can view the breeding value as the portion of the parent’s genotype that causes their offspring’s phenotype to deviate from the population mean (the connection to V_A should be clear since V_A determines how the population mean can shift between generations). More simply, the variance in individual breeding values *is* the additive genetic variance (V_A).

4 Fitness, Selection, and Adaptation

In section 2 we have talked a lot about selection without formally defining it. In order to introduce selection more formally, we need to examine an important factor in the definition of selection: fitness. Fitness is a central but refractory concept in evolutionary biology. This is because researchers think about fitness in different ways depending on the question at hand. Behavioral ecologists, demographers, and game-theoreticians all use different measures of fitness that operate on different time-scales and contain different assumptions. While all conceptualizations of fitness concern reproduction, there are some key differences as to how fitness is best measured. One of the most basic distinctions among measures of fitness used by these different researchers is that of “rate-insensitive fitness” and “rate-sensitive fitness.” Rate-sensitive concepts of fitness measure evolutionary success on a per-generation basis, whereas rate-insensitive concepts of fitness measure success in absolute time. A very common rate-sensitive measure of fitness is net reproductive rate (R_o). This measure is particularly useful to behavioral ecologists because it ties into lifetime reproductive success. That is, if one is looking at a bunch of animals in a population, the average of their lifetime reproductive success (LRS) is R_o . Thus estimating lifetime reproductive success in individuals allows one to estimate the per-generation rate of offspring production. In practice, calculating LRS takes a lot of effort and thus researchers often resort to calculating components of LRS. Thus when one measures survival, mating success, or rate of offspring produced over a given period, these measures constitute components of LRS. Most behavioral ecological applications of equation 16 use LRS or some components thereof. The advantages of this approach is that it often allows one to define the appropriate component of LRS in order to address a particular question about the type of selection (e.g., viability, intrasexual, fecundity, intersexual) acting on traits. For example, if one is interested in sexual selection, it may be appropriate to use “male mating success per mating season” as the measure of fitness. In this case, traits that covary with this fitness definition, as specified by the beta coefficients in equation 16, measure the direct action of selection on these traits due to male “mate-getting” ability. Likewise, defining fitness as lifetime survival (or survival over a given period) will allow one identify the traits associated with viability selection. When components of LRS are used to estimate selection (e.g., survival), they estimate selection over a single episode of selection, not over a lifetime. Arnold and Wade [8] and Wade and Kalisz [93] discuss the complications and potential solutions to this problem. Conner [35] and Conner and Hartl [37] provide a lucid description of these methods and Brodie and Janzen [16] provide some caveats.

LRS, as discussed above, is a simple tally of offspring produced over a lifetime, but it doesn’t tell us *when* offspring are produced during a lifetime nor does it take into account whether the population is growing or shrinking. Because selection can act strongly at some ages and less strongly at other ages, it is better to think of fitness in terms of the entire life cycle of the organism, from birth to death. It is necessary to do this because it matters *when* individuals reproduce during their lifetime; for example, in an expanding population it pays to reproduce earlier than your conspecifics. Given this, we need to find a fitness measure that accounts for the timing of reproduction as well as changes in population size, since changes

in the population growth rate can influence the strength of selection acting on different age classes [19]. That is, one needs a fitness measure that is insensitive to generation times, or “rate-insensitive.” A fitness measure that satisfies these criteria is the intrinsic rate of increase of a population (or population growth rate), often denoted as λ or r . These two measures are connected by $r = \log_e \lambda$. [This measure of fitness also connects to the concept of mean fitness in population genetics, usually denoted as \bar{w} , in that $r \approx \log_e \bar{w}$, assuming that the per generation growth rate is not too large]. The population growth rate, λ , is a demographic measure of the *eventual* rate of population growth, assuming the environment does not change. We stress *eventual* because it measures the rate of growth of a population in the long-term, or once the proportion of individuals in different age classes remain stable. The population growth rate, λ , is calculated from the average rates of survival, growth, and fertility in the population, which are embodied in the projection matrix, \mathbf{A}_t , described in section 6.3. Populations grow in size when $\lambda > 1$, remain the same size when $\lambda = 1$, and shrink in size when $\lambda < 1$. Since λ takes into account timing of offspring production it is often considered a superior measure of fitness than R_o . Thus to a demographer, λ is the proper measure of fitness and R_o is only a component of fitness. The drawback of using λ is that it is tricky to estimate on individuals, since λ is calculated from averages—hence, the preference for R_o among behavioral ecologists. There have been attempts to develop methods that use a rate-insensitive measure of fitness that can be estimated from individuals [66] [38] [18].

For completeness, we discuss another measure of fitness used mostly by theoreticians who study evolutionary dynamics. This measure of fitness pertains to *invasibility* (also called the invasion exponent). Consider a population in which individuals have the phenotype “have one offspring in year t , then die” (this is their “strategy”). Now consider a mutation that occurs in one individual that codes for a new strategy, such as “have one offspring in year $t + 1$ and one in year $t + 2$, then die”. Game theoreticians want to determine if this mutant strategy can increase in frequency in the population in the face of the existing strategy. To do this, they calculate the growth rate of the existing strategy (λ_e) and the growth rate of the mutant strategy (λ_m). Fitness, then, is viewed as the ability (or lack of ability) of a mutant strategy to invade (i.e., increase in frequency) a population predominated by different strategy. It is measured as the difference in growth rates between the two strategies ($\lambda_e - \lambda_m$). An evolutionary stable strategy is a population containing a single strategy that cannot be invaded by a mutant strategy.

Invasibility is considered the master fitness concept, but it is rarely used in primatology. On the other hand, both λ and R_o have been empirically estimated on wild primate populations. In short, λ is probably a better measure of fitness than R_o but λ is difficult to operationalize since it is a property of a population and is difficult to apply to short time intervals; LRS is easy to measure on individuals animals, operates over short-time intervals, but does not take into account the timing of reproduction. The two measures of fitness R_o and λ do give equal results when $\lambda = 1$. Technical but informative discussion of fitness are found in Charlesworth [23], Benton and Grant [9], Brommer [17], and Caswell [22]. What should be clear about all three measures of fitness, however, is that they all tie into our notion of fitness-as-reproductive success, independent of the time-scale or timing involved.

We now turn to the other major factor in evolutionary change: selection. Intuitively, we know that selection can change phenotypic variation in a population in non-random ways. Here, we want to define selection more rigorously. The general manner in which selection can change phenotypes is by acting on the mean, the variance, or the joint variation of a bunch of traits in a population. Thus, if we are looking at height, selection can act to increase or decrease the mean height (this is positive and negative directional selection, respectively), to increase or decrease the variance in height (this is disruptive and stabilizing selection, respectively) or to act on height along with another trait, such as weight (this is correlational selection). Correlational selection refers to the fact that combinations of traits interact to produce a fitness effect. Since we know that selection acts on traits with respect to the fitness differences those traits confer, we can visually depict the action of selection by plotting the relationship between fitness and phenotype. More formally, what we are trying to visualize is how variation in fitness covaries with phenotypic variation (e.g., equation 3). In fact, one can rigorously define selection as *the covariance between fitness and those phenotypes that causally influence fitness differences*. We need to emphasize the causality part because drift and/or other evolutionary forces can also produce a non-random relationship between fitness and phenotype. Thus, measuring selection is not simply a matter of regressing fitness on phenotype but understanding the causal relationship between particular trait values and how they determine variation in fitness [80].

That said, selection can be estimated using linear and non-linear least-squares regression models. In principle, any measure of fitness described above can be used; however, in practice, usually researchers regress some measure of LRS against phenotype. When regressing fitness on phenotype, the regression coefficients tell us something about the strength (i.e., magnitude) and type (directional, stabilizing, etc.) of selection acting on the phenotype. This is because regression coefficients themselves are defined as $\text{Cov}(w, z)/\text{Var}(w)$ —if w is fitness and z is phenotype, then it is clear that the regression coefficient captures something about the covariance between fitness and phenotype. In fact, in equation 15, the β coefficients are partial regression coefficients of fitness and phenotype [60]. In a multivariate framework, the partial least squares regression coefficients (β) reveal how much a particular trait “predicts” fitness when holding the other traits constant. Figure 1 shows examples of types of selection as well as the expected response to selection, assuming that the traits under investigation possess heritable variation. The response to selection shows what aspects of the phenotype distribution (i.e., mean, variance, covariance) change given a particular form of selection. Also shown is the regression equation used to estimate the coefficients describing different types of selection. Thus it is not only possible to estimate directional selection (β) but also stabilizing/disruptive selection (γ_{ii}) and correlational selection (γ_{ij}). The γ_{ii} term captures changes in the variance of trait i , whereas the γ_{ij} term captures changes in the covariance between traits i and j (Figure 1). It is important to note that selection must act on that component of phenotypic variation that is heritable (i.e., the breeding value—see Section 3) in order to produce an evolutionary response. It is possible that a non-zero correlation between fitness and phenotype could occur in which selection is acting on that portion of phenotypic variation that is determined by the environment [77]. For example, a positive covariance between survivorship and body mass might be found in humans, but all the variation in body mass is due environmental variation (e.g., some individuals have

deficient diets). Mitchell-Olds and Shaw [67], among others (e.g., [64]) discuss the statistical caveats in applying these methods to field data.

Our discussion of selection and fitness leads to another important but contentious topic in evolutionary biology—adaptation. For the most part, adaptation is viewed from two different angles. The first conceptualization concerns *etiology*, where priority is given to either the initial selection pressures that produced the adaptation or the derived status of a trait (produced by selection) within a clade. This definition is adopted in slightly different forms by Gould and Vrba [48]; Coddington [34], and Harvey and Pagel [49], among others. Usually under this definition adaptations are identified using the comparative method and/or design principles. The second conceptualization of adaptation is that of *current utility*, where priority is given to the immediate fitness benefits that the trait confers independent of its etiology. Reeve and Sherman [79], Fisher [42], and Bock [14] have all proposed definitions of adaptation that pertain to current utility. Usually under this definition adaptations are identified by measuring selection or developing optimality models. With its emphasis on immediate fitness benefits and trait values, equation 16 falls into a natural allegiance with the current utility approach [5] [53]. That is, equation 16, is able to get at which traits might qualify as adaptations because such traits can be identified as the direct targets of selection. To the extent that there is a non-random and causal association between trait values and fitness values, one can conclude that such trait qualifies as a potential adaptation. In this regard, equation 16 provides insight into the adaptive process. It allows one to answer questions such as: Which traits are under direct selection? Which traits evolve due to a correlated response? Answering such questions can point to the ecological circumstances that produce particular selection pressures and why certain trait values confer higher fitness. Naturally, additional evidence should be brought to bear on the issue; for example one can further test the adaptive basis of a trait using the comparative method or experimental manipulation.

5 Model assumptions and the role of models in science

Equation 16 is a model of the evolutionary process. In this section we discuss the various assumptions that are built into equation 16. It is likely that most of these assumptions are often violated in nature. Consequently, our discussion will turn to the role of models in science. When applying equation 16 across generations, the following assumptions are made:

1. The traits under study are multivariate normal.
2. The \mathbf{G} matrix remains constant (which implicitly suggests that allele frequencies do not change much due to selection or drift).
3. The additive genetic and phenotypic variances change more slowly than the average trait values.

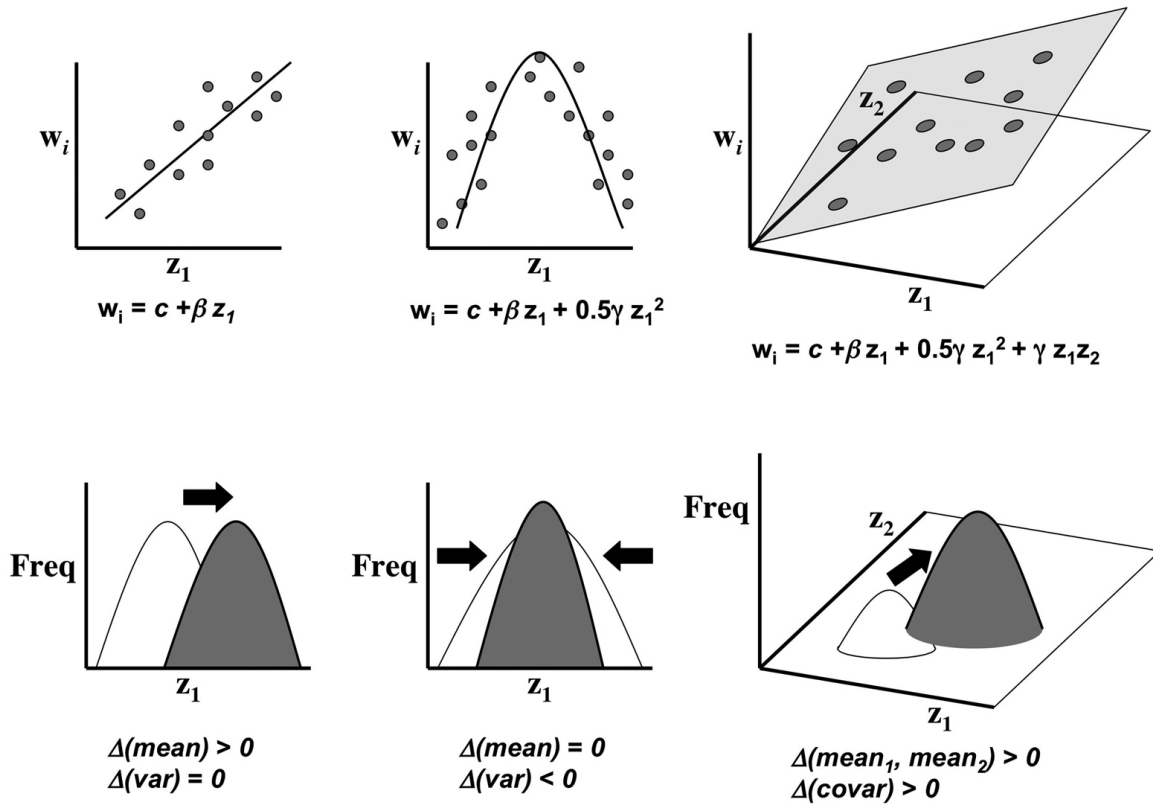


Figure 1: A graphic relationship between fitness variation and phenotypic variation and equations describing this relationship (top). The response to selection is shown below and is a product of the strength of selection + heritable variation in the trait. Regarding correlational selection, the shape of the fitness surface depends on the various selection coefficients; therefore, the fitness surface will not always be flat, as depicted here

4. No genotype by environment interaction.
5. The environment remains constant from generation to generation.
6. No maternal/paternal/indirect effects (for example, an offspring’s phenotype is not partly a function of its mother’s environment, or the social environment provided by conspecifics).
7. No epistatic effects among loci contributing to the traits under study.

Ideally, if the above assumptions are not violated, then one could iterate equation 16 over many generations and predict the trajectory of mean trait values through “phenotype space”. In reality, it is likely the case that the environment changes, as do the terms in the \mathbf{G} matrix. Regarding the latter, one can conclude that the \mathbf{G} matrix itself evolves, and this topic is an area of active research [84, 74]. In most ecological applications of equation 16, however, the researcher is interested in the strength of selection and/or the additive genetic variances of the traits, not their long-term trajectories [82]. In this case, there is little need to understand the long-term changes in the \mathbf{G} matrix, since the β coefficients and trait heritabilities provide insight into immediate adaptive processes.

With so many assumptions going into equation 16, one might ask what is the usefulness of constructing such an idealized equation in the first place. To begin this discussion we make the blatant assertion that all models make assumptions and thus *all models are wrong to some degree*. Given this statement, why bother developing a “wrong” model. For starters, models are constructed not for their truthfulness, but for their heuristic ability to depict a certain property of a system under idealized conditions. More colloquially, models ask, “How much of the real world can we throw away and still get a good answer?” Very often, the answer shows what is possible when all but one or two key factors are left out. To the extent that the model predicts and/or provides a good description of what we see out in nature, then we have a good model. This is because our model has shown us which factors are important given our observations. Caswell [20] makes the apt analogy: models are to theoretical problems as experiments are to empirical problems. Good experiments do not include several factors at once and the same is true of models. Our understanding of causality would be clouded if we included ten different factors in a single model or experiment—how could we isolate the key factor that “drives the system”, so to speak? It is much more logical to ask what key factors appear to be the most important with respect to the question at hand. From this one can develop a simplified characterization of a particular problem in order to see how it stands up against reality (e.g., how traits change under directional selection).

Researchers who are unfamiliar with modeling techniques might be tempted to criticize a model for what it has excluded. It is always easy to suggest what a particular model has left out; it is more important to understand if the excluded material affects one’s conclusions given the question at hand. And the only way to do so is to build a model and see what it predicts in the absence of such factor. By simplifying and isolating the key properties

of the system under study, a properly constructed model allows the researcher to control, and incrementally add, which factors are likely important. Equation 16 distills evolutionary change due to directional selection down to two key factors— β (selection) and \mathbf{G} (genetic variation). It allows us, quite accurately, to predict how mean trait values will change over short-time intervals. No doubt other factors are also relevant but we wouldn't necessarily know that β (selection) and V_A (genetic variation) are important if our model had included various factors at the outset.

6 Extensions to specific evolutionary phenomena

In this section, we show how equation 16 can be modified and applied to particular evolutionary phenomena such as sexual selection, life history theory, allometry and constraints, ecological morphology, and social evolution. As will become evident, addressing such topics often involves nothing more than carefully redefining one or more terms in equation 16 such that they capture the basic properties of the system under study.

6.1 Sexual selection and mate choice

Here we show how equation 16 can be modified to capture the dynamics of phenotypic evolution due to variation in mate acquisition. We focus on two major topics: models of female mate choice and the evolution of sexual size dimorphism. Darwin's original conceptualization of sexual selection was predicated on explaining the existence of male traits that were seemingly detrimental to viability (e.g., a conspicuous peacock's tail) or had no intuitive functional purpose (e.g., extremely large antlers); his solution was to suggest that some male traits that were bad for survival were good for obtaining mates. A major component of sexual selection concerns female mate choice. Males displaying elaborate traits were more likely to be chosen by females for mating independent of how their trait influences their viability. In this regard, modeling sexual selection of female preferences requires that we consider three traits—a trait denoting female preference (p), a male display trait (t), and a trait pertaining to viability (v)—and two types of selection: sexual selection (due to variation female preferences), and non-sexual selection (which captures selection that operates on aspects of fitness not due to mate acquisition). We can modify equation 16 to capture female preference models of sexual selection. This approach is developed and treated in detail by Fuller et al., [44]. Because two types of selection are present, the selection gradient is broken up into two components, one due to non-sexual selection (β_n) and one due to sexual selection (β_s). The \mathbf{G} matrix contains the variances and covariances among the three traits. The general model is

$$\Delta \bar{\mathbf{z}} = \mathbf{G}(\beta_n + \beta_s) \tag{21}$$

and expanding the terms

$$\begin{pmatrix} \Delta \bar{z}_t \\ \Delta \bar{z}_p \\ \Delta \bar{z}_v \end{pmatrix} = \begin{pmatrix} V_A(t, t) & C_A(t, p) & C_A(t, v) \\ C_A(t, p) & V_A(p, p) & C_A(p, v) \\ C_A(t, v) & C_A(p, v) & V_A(v, v) \end{pmatrix} \times \left(\begin{bmatrix} \beta_{n,t} \\ \beta_{n,p} \\ \beta_{n,v} \end{bmatrix} + \begin{bmatrix} \beta_{s,t} \\ \beta_{s,p} \\ \beta_{s,v} \end{bmatrix} \right) \quad (22)$$

Female preference models of sexual selection come in many different forms, for example Fisher’s runaway selection model, the good genes model, sensory bias models, and direct benefit models. As shown by Fuller et al. [44], the predictions from these different models specify which terms in the \mathbf{G} matrix and selection gradient vectors should be important. Many of these sexual selection models require that the selection coefficients and variances and covariances change as the system evolves. Thus, Fuller et al., showed the initial and equilibrium conditions for each model. Below, we discuss two models that are likely to be important in primatology: the “good genes” model and the “direct benefits” model.

In the good genes model, males are chosen by females because males possess some traits that confer high viability, which the female’s offspring inherit. This model was developed to explain why females choose particular mates despite the fact that their mates offer little in terms of paternal care (i.e., females are choosing mates based on male’s genes, not their paternal ability). Good genes models possibly operate in mandrills [86]. Two assumptions of the good genes models are that the cost of the male trait (t) increases with the size of the trait and that the cost of the female preference (p) increases with females choosiness. Both are realistic: a male trait may be an elaborate display (e.g., some African antelopes “stot” or jump up and down in front of a predator in what is interpreted as a display of vigor) and a lengthy display may attract predators; equally, females who spend too much time looking for the right mate may spend less time feeding or may be more vulnerable to predation. Under the good genes model, at equilibrium, the important terms in equation 18 are highlighted below

$$\begin{pmatrix} \Delta \bar{z}_t \\ \Delta \bar{z}_p \\ \Delta \bar{z}_v \end{pmatrix} = \begin{pmatrix} V_A(t, t) > 0 & C_A(t, p) > 0 & C_A(t, v) > 0 \\ & V_A(p, p) > 0 & \mathbf{C}_A(\mathbf{p}, \mathbf{v}) > \mathbf{0} \\ & & V_A(v, v) > 0 \end{pmatrix} \times \left(\begin{bmatrix} \beta_{n,t} < 0 \\ \beta_{n,p} < \mathbf{0} \\ \beta_{n,v} > 0 \end{bmatrix} + \begin{bmatrix} \beta_{s,t} > 0 \\ \beta_{s,p} \equiv \mathbf{0} \\ \beta_{s,v} > 0 \end{bmatrix} \right) \quad (23)$$

In equation 23, the negative $\beta_{n,t}, \beta_{n,p}$ coefficients capture the action of non-sexual selection on the high costs of having an elaborate trait as well as being too choosy. The positive $\beta_{n,v}, \beta_{s,p}, \beta_{s,t}$ terms reflect the fact that viability (v) is positively selected for (since males survive in spite of the elaborate trait), and that both the trait and the viability it confers are under positive sexual selection. In this model, there is never any direct sexual selection for preference (denoted by $\beta \equiv 0$). The non-sexual selection against preference ($\beta_{n,p} < 0$) is counteracted by the indirect positive selection for female preference due to the positive genetic covariance term $C_A(p, v)$; thus females are selected to prefer elaborate males, but this preference is a correlated response due to positive sexual and nonsexual selection for viability ($\beta_{n,v} > 0, \beta_{s,v} > 0$).

In the direct benefits model of mate choice, females choose males because of the immediate qualities they can confer to the female and her offspring (such qualities include things like

paternal care or protection from predators). In this model, there is no trait—and hence no selection—that signals good genes to the female (e.g., a trait that signals increased viability). Thus at equilibrium the direct benefits model is,

$$\begin{pmatrix} \Delta \bar{z}_t \\ \Delta \bar{z}_p \\ \Delta \bar{z}_v \end{pmatrix} = \begin{pmatrix} V_A(t, t) > 0 & C_A(t, p) & C_A(t, v) \\ & V_A(p, p) > 0 & C_A(p, v) \\ & & V_A(v, v) \end{pmatrix} \times \left(\begin{bmatrix} \beta_{n,t} < 0 \\ \boldsymbol{\beta}_{n,p} < \mathbf{0} \\ \beta_{n,v} \end{bmatrix} + \begin{bmatrix} \beta_{s,t} > 0 \\ \boldsymbol{\beta}_{s,p} > \mathbf{0} \\ \beta_{s,v} \end{bmatrix} \right) \quad (24)$$

In this model, the important terms are the following: non-sexual selection against preference (females incur a survival cost for being too choosy, $\beta_{n,p} < 0$), and sexual selection for preference (females receive direct benefits for selecting particular mates, $\beta_{s,p} > 0$). There is no additive genetic covariance between preference and viability.

These are just two of the many models of mate choice. The contribution from Fuller et al., was to put these models into a common framework; this allows for the easy identification of the key terms that distinguish each model. As Fuller et al., point out estimating such terms can be daunting but doing so will allow better resolution to the processes shaping mate choice in wild populations. Estimating some of the terms in equations 23 and 24 would go a long way toward sorting out whether female primates gain direct and/or indirect benefits when choosing mates (reviewed in [71]). We should note that our treatment above is highly simplified; for example, we have ignored mutational input—a pivotal term that can change the predictions of a particular model (compare [54] with [76]). Our goal here was to illustrate the basic quantitative genetic framework of sexual selection via mate choice. Kokko et al., [55] provides additional commentary and analysis on these models and overall framework.

6.2 Sexual selection and size dimorphism

We now turn to the evolution of sexual size dimorphism (SSD). In many primate species, SSD likely evolved due to male-male mating competition [75]. The first explicit quantitative genetic treatment of SSD was provided by Lande [58]. Our explanation follows that given in [78]. To set up his treatment, Lande introduced a model of stabilizing selection. This type of selection can be captured mathematically by denoting $\boldsymbol{\theta}$ as a vector of optimum trait values and $\bar{\mathbf{z}}$ as the mean values of the traits in question, then one can represent stabilizing selection as

$$\boldsymbol{\theta} - \bar{\mathbf{z}}$$

Thus when each trait is at its optimum, the expression goes to zero. This makes sense, in that if we are modeling how traits change under stabilizing selection, we would expect no more change due to selection once traits reached their optimum value, (i.e., $\Delta \bar{\mathbf{z}} = 0$). Lande's full model of stabilizing selection was thus

$$\Delta \bar{\mathbf{z}} = \mathbf{G}(\boldsymbol{\omega} + \mathbf{P})^{-1}(\boldsymbol{\theta} - \bar{\mathbf{z}}) \quad (25)$$

\mathbf{P} and \mathbf{G} are the phenotypic and additive genetic variance-covariance matrices respectively. The new terms in this equation are a matrix of stabilizing selection values ($\boldsymbol{\omega}$) that specify the strength of stabilizing selection (along the diagonal) and correlational selection (on the off-diagonal). If there are n traits in the model, then $\boldsymbol{\omega}$ is a $n \times n$ matrix. If the entries along the diagonal of $\boldsymbol{\omega}$ are much greater than the corresponding values in the \mathbf{P} , then selection is weak, whereas if they are much smaller than the phenotypic variance then selection is strong. To model the effects of selection acting separately on males and females, Lande added more terms to equation 25. The first term is a vector of sexual selection coefficients ($\boldsymbol{\beta}_s$) that specify the strength of directional selection acting on either males or females, the second term was a modified additive genetic variance-covariance matrix that specifies the amount of genetic covariance between males and females (\mathbf{B}). Given these new terms, Lande measured how male traits change as a result of directional sexual selection in males using the following equation (an analogous equation was developed for females but we don't show it here, see [78]):

$$\begin{aligned} \Delta \bar{\mathbf{z}}_m &= 0.5 \mathbf{G}_m (\boldsymbol{\omega}_m + \mathbf{P}_m)^{-1} ((\boldsymbol{\theta}_m - \bar{\mathbf{z}}_m) + \boldsymbol{\beta}_{s,m}) \\ &+ 0.5 \mathbf{B} (\boldsymbol{\omega}_f + \mathbf{P}_f)^{-1} ((\boldsymbol{\theta}_f - \bar{\mathbf{z}}_f) + \boldsymbol{\beta}_{s,f}) \end{aligned} \quad (26)$$

This equation, while seemingly daunting, has three main parts: 1) stabilizing selection acting on male traits (captured by the $\boldsymbol{\omega}_m, \mathbf{P}_m, \boldsymbol{\theta}_m, \bar{\mathbf{z}}_m$ terms); 2) directional sexual selection acting on male traits (the $\boldsymbol{\beta}_{s,m}$ terms); and 3) the effect of stabilizing and sexual selection acting on female traits (the same terms in parts 1 and 2 but subscripted with f). The two additive genetic variance-covariance matrices specify the covariances in traits within a male (\mathbf{G}_m) and the covariances between males and females (\mathbf{B}).

The evolution of sexual dimorphism depends largely on the size of the genetic covariance between the sexes (\mathbf{B}) as well as the pattern of selection on males and females. The evolution of sexual dimorphism using equation 26 has largely been studied through simulation. In this case, the researcher specifies initial trait values for males and females as well as the strength of the covariances and the strength of stabilizing and sexual selection. A simulation is run using equation 26, with males and females initially starting out at the same size (this is akin to stating that natural selection favors males and females at the same optimal size). One can then plot the trajectory of changes in male and female body size against time. When directional sexual selection is acting on male body size but not female size, most simulations show a characteristic trajectory of body size values for males and females, as discussed in [40] (Figure 2). Initially, both male and female size increases due to direct sexual selection on male body size with the increase in female body size resulting from a correlated response (this period of monomorphism can last numerous generations depending on the magnitude of selection and genetic correlations). After this initial and equal increase, however, the mean male and female trait values slowly begin to diverge, and the rate of divergence depends on the magnitude of the terms in \mathbf{B} matrix. Over many generations, the two sizes settle on their new optimal sizes: Females return to the size favored by natural selection, whereas male size is a balance between the effects of natural selection and sexual selection. It is important to note that the above model is highly simplified. Fairbairn [40] and Reeve and Fairbairn [78] present good discussions of the assumptions and limitations of these models. A recent and

interesting alternative model for SSD as it applies to primates is given in [47].

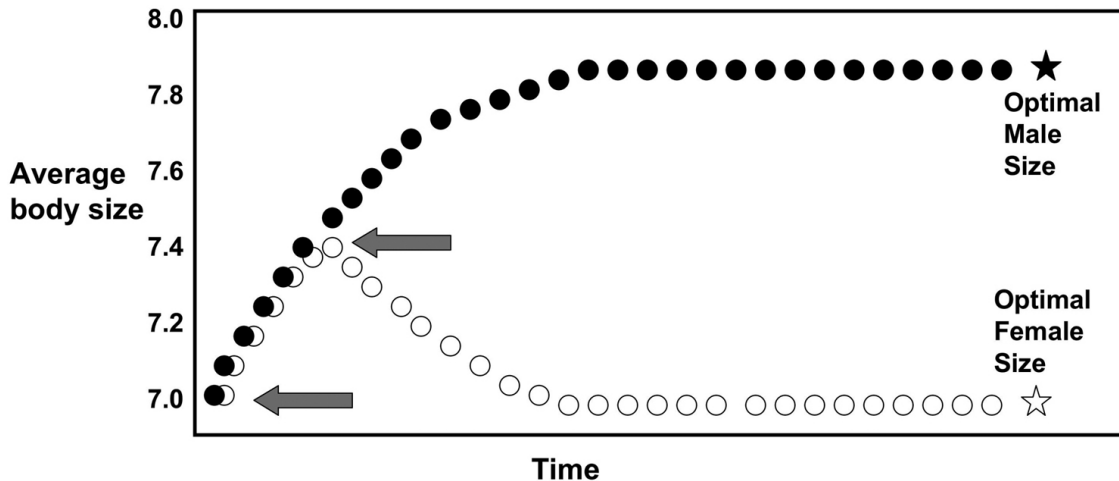


Figure 2: Schematic of a typical simulation of equation 26 for the evolution of sexual size dimorphism (redrawn from [40]). Sexual selection operates on male body size and natural selection selection operates on male and female body size in the form of stabilizing selection. The actual trajectory of body sizes depends on the parameter values in eq. 26 and population size; the dots do not represent single generations, they simply show the basic trajectory. Note that monomorphism occurs during the phase bracketed by arrows. See text for discussion.

Such simulations have interesting implications for the evolution of sexual size dimorphism in primates. As shown in Figure 2, there is an initial period in which males and females are monomorphic, even though both sexes are increasing in size (the portion of the trajectory between the arrows). This suggests that in monomorphic primates (e.g., *Callicebus* or some *Propithecus* spp.), similar size may be due to either no sexual selection acting on males (likely in the case of *Callicebus*), or that equal body size is due to non-equilibrium conditions, where equal male and female body size is a transient condition, as has been proposed in some diurnal lemurs [92] (but see [61]).

6.3 Life history theory

Imagine we are interested in understanding how selection acts on the life history traits of a baboon. Recall that life history traits are the major phenotypic manifestations of fitness itself. In this regard, all we are doing is recognizing that things like age at sexual maturity, survivorship, and age at death are phenotypic traits and thus can be targets of selection similar to morphological traits. The best way to understand how selection acts on such traits is to develop a life cycle model. The life cycle model (or graph) should capture the particular life history traits one wants to examine. An example of a life cycle is given in Figure 3. This life cycle has four stages corresponding to 1yr olds, juveniles, subadults/adults without offspring, and mothers with infant. The arrows on the life cycle specify which

transitions are possible from year to year (i.e., you can go from yearling to juvenile but not vice versa) and the coefficients under the arrows give the average probability of making the transition specified by the arrows each year as well as the average probability of giving birth (transition F_4). The coefficients can also be represented in a matrix form and such a matrix is called a projection matrix, by convention denoted as \mathbf{A}_t , with entries a_{ij} denoting the i th row and j th element in matrix \mathbf{A} (to keep with this convention, we use a_{ij} to denote life history traits, rather than z in this section).

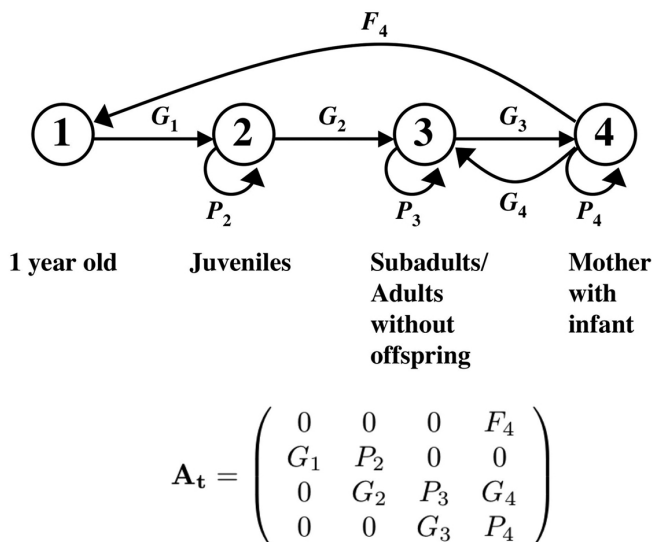


Figure 3: Hypothetical life cycle (top) and projection matrix (bottom) for a baboon. This life cycle contains four stages. The coefficients on the arrows represent the probability of moving within (P_i) and among (G_i) stages each year. The F_4 term represents the probability of moving into the maternity stage (i.e., giving birth) multiplied by the probability that the infant will survive to stage 1. These probabilities enter into the projection matrix (\mathbf{A}_t). The dominant eigenvalue of the projection matrix provides the population growth rate, λ (see section 4). Sensitivities of λ to entries in the projection matrix can be interpreted as directional selection gradients, as discussed in the text.

It should be apparent that the coefficients tell us something about the life history of the organism because they specify probabilities of survival, growth, and fertility. [Obviously it is possible to develop many alternative life cycles for a baboon, including those using ages instead of stages—the key is to develop a life cycle that represents the biological question you’re interested in.] What we want to find out here, is how directional selection acts on these coefficients. Answering this question will provide insight into which stages are important determinants of fitness. In 1982, Lande demonstrated how to connect the demographic aspects of life history theory to quantitative genetics [59]. What he was able to show was that response to selection on life history traits could be calculated as

$$\Delta \bar{\mathbf{a}} = \frac{1}{\lambda} \mathbf{G} \nabla \lambda \tag{27}$$

where,

$$\nabla\lambda = \begin{pmatrix} \frac{\partial\lambda}{\partial a_{11}} \\ \frac{\partial\lambda}{\partial a_{12}} \\ \vdots \\ \frac{\partial\lambda}{\partial a_{ij}} \end{pmatrix}$$

In words, this equation reads: the change in the mean phenotype of a vector of traits is equal to the reciprocal of mean population growth rate (λ) multiplied by the \mathbf{G} matrix multiplied by the selection gradient. The biological interpretation of the terms in the equation is as follows. The $\Delta\bar{\mathbf{a}}$ is the response to selection, which captures the following process: the change per unit time in average phenotype is equal to the difference between the mean breeding value at birth and at age x , divided by the generation time. This formulation encapsulates the idea that at birth the genotypic and phenotypic probabilities of survival, growth, and fertility are similar, but as the population ages the phenotype distribution changes because some individuals fail to survive or breed. What Lande derived was the fact that as phenotypes change through time, the change in phenotypes is due to differences in survival among particular genotypes (from birth to age x), weighted by the number of offspring each genotype produced. On the right hand side of the equation, the \mathbf{G} , is the G-matrix—a matrix containing the additive genetic variances and covariances for the life history traits. The last term is a vector of selection gradients. Formally, this is an array of partial derivatives of mean population growth rate with respect to each trait. Such partial derivatives are analogous to the β coefficients in equation 15 above (that is, $\beta \approx \frac{\partial\lambda}{\partial a_{ij}}$). Like beta coefficients, the partial derivatives measure the direct action of selection acting on each life history trait while all others are held constant.

Lande’s final equation (equation 27) is very similar to equation 16. What is underappreciated is the elaborate mathematics used to merge the demographic aspects of life history theory, as captured by the life cycle, and the basic framework of quantitative genetics, as expressed in equation 16. The clearest exposition of Lande’s derivation is given in Caswell [21]. In connecting the two, it was necessary to make adjustments to equation 16 to incorporate age or stage structure and overlapping generations while making sure that the measure of fitness relates to the transitions in the life cycle graph. Further, as in other conceptualizations of multivariate evolution, the equation must account for correlations between traits, since life history theory is predicated on idea that many traits have negative correlations (i.e., trade-offs). Finally, Lande needed to make sure the response to selection is measured in absolute time, not per generation. Lande’s equation provides a critical link between understanding life cycle evolution and quantitative genetics. There is a substantial body of theory for analyzing projection matrices and an important concept in this theory is the concept of sensitivity [22]. In the most common usage, sensitivity measures the degree to which population growth rate would change given some change (i.e., perturbation) in the projection matrix entry, holding all other entries constant (recall that the matrix entries give us the probabilities of engaging in some sort of life history transition). Mathematically, sensitivities are represented as $\frac{\partial\lambda}{\partial a_{ij}}$. Immediately, this should look familiar: $\frac{\partial\lambda}{\partial a_{ij}}$ is a component of the selection gradient in Lande’s equation (see equation 27, above). Thus, sensitivities in matrix population models

have an interpretation as selection gradients, this is because they measure how fitness (as represented by population growth rate, λ), depends on (or “covaries with”) a particular life cycle transition, a_{ij} , and we learned in Section 4 that selection can be formally defined as the covariation between fitness and phenotype. Sensitivities are a key part of demographic theory; Lande’s contribution was to anchor them into an evolutionary quantitative genetic framework. We should note that such selection gradients ($\frac{\partial \lambda}{\partial a_{ij}}$) measure directional selection; to measure stabilizing/disruptive selection it is necessary to take the second derivatives of fitness with respect to each trait (i.e., $(\frac{\partial^2 \lambda}{\partial a_{ij}^2})$). There are several programs that enable the calculation of sensitivities from projection matrices (e.g., RAMAS, Vortex) and they can also be calculated using a few lines of Matlab code [22] or R code [51]. Upon their calculation, one can interpret the sensitivity value as the magnitude of directional selection acting on that transition in the life cycle (see Section 7.1).

6.4 Evolutionary constraints

In this section and the next one we address topics that examine how traits interact during their evolution. We will consider the \mathbf{G} matrix first. The \mathbf{G} matrix tells us, in part, how two or more traits will respond to selection based on their genetic covariances. Depending on the sign of these covariances, the response to selection (i.e., the change in mean trait values) may not be in the direction of highest fitness. Here, we define “constraint” as a bias in the selection response that prevents a population from reaching its maximum fitness (maximum fitness could be determined a priori, for example, through an optimality model). For example, assume that height and weight have a negative genetic covariance. This would suggest that direct selection to increase height would indirectly decrease weight (and vice versa). Now imagine an ecological scenario in which both increased height and increased weight conferred the highest fitness (for example, tall and heavy individuals may have an advantage during aggressive contests). Consider, that there is direct positive selection to increase height and direct positive selection to increase weight. Because of the negative genetic covariance between height and weight, the response to selection will not move in the direction of *both* increased height and weight. Thus, the negative genetic covariance constrains, or biases, the selection response away from the trait values which confer the highest fitness [26]. The \mathbf{G} matrix can also enhance, rather than bias, the selection response. In this case, if two traits are positively correlated and there is direct positive selection on both traits, the selection response will be enhanced. Conner and Hartl [37] provide a nice table illustrating how evolutionary change will be enhanced or constrained depending on the sign of the genetic covariance and the whether directional selection is acting to increase or decrease the mean value of a trait (Table 2).

It is important to note that a negative genetic covariance can exist despite a positive phenotypic covariance among traits. Cheverud [27] suggests that often phenotypic covariances can be used as a proxy for genetic covariances, when the latter cannot be measured directly. However, this proxy appears to hold better for morphological traits, but not life history

Table 2: Ways in which genetic covariances enhance or constrain the selection response depending the pattern of selection acting on each trait (from [37])

Sign of genetic covariance:	Sign of β for each trait:	
		Positive
	<i>The selection response will be:</i>	
Positive	Enhanced	Constrained
Negative	Constrained	Enhanced

traits [82]. In fact, life history theory suggests that there should be trade-offs among traits such as survival and reproduction, as well as current and future reproduction [88] [83] (see Section 7.2). Here, the \mathbf{P} and \mathbf{G} matrices are expected to contain negative covariances between traits; such covariances can bias the response to selection toward non-optimal trait combinations. Arnold [7] provides a good discussion of evolutionary constraints and Pigliucci [72] provides a recent review and critique of the “ \mathbf{G} matrix concepts” in evolutionary biology. Two additional points are worth mentioning about the \mathbf{G} matrix. The first is that such constraints and biases in the selection response operate over the short-term—predicting long-term trajectories requires the major assumption that the terms in the \mathbf{G} matrix do not change. Second, the \mathbf{G} matrix tells us how the *mean* values of traits respond to selection due to a common genetic control. In principle, one can have associations between trait distributions that involve other moments of the distribution (e.g., variances, skewness, and kurtosis). For example, selection on the mean of one trait can change the variance or skewness of another trait. Rice [80] [81] develops a body of theory for the evolution of *entangled traits*; traits were selection on the n^{th} moment of one trait’s distribution changes the m^{th} moment of another trait’s distribution due to developmental associations among traits [81].

6.5 Allometry

Allometry is the study of scaling relationships among anatomical systems. Most often, allometry is studied in the context of how one morphological measurement changes in size or shape with respect to changes in another morphological measurement. A lot of work in allometric studies seeks a biological interpretation for the scaling relationship among two or more traits. Because allometry involves two (or more) characters that covary, it suggests that it can be studied using phenotypic or genetic variance-covariance matrices. In fact, the original derivation of equation 16 above, was used in the context to study brain-body mass allometry in mammals by Lande [57]. What was known at the time was that brain-body mass had an allometric coefficient of 0.2-0.4 among closely related species (i.e., within the same genus). One of the goals of Lande’s analysis was to figure out if this allometric scaling relationship could be explained by directional selection on body size

with brain size being a correlated response. Here, we'll look at the ways in which equation 16 can be used to gain insight into allometric relationships. We'll denote z_1 as the log of body size and z_2 as the log of brain size. Further, for notational purposes, we will define additive genetic variances and covariances for traits as $\sigma^2(A_1, A_1)$ and $\sigma(A_1, A_2)$, respectively (i.e., $C_A(1, 2) = \sigma(A_1, A_2)$ from above); we change the notation since it is easier to keep track of how a variance, $\sigma^2(A_1, A_1)$, can be converted to a standard deviation by taking its square root, $\sqrt{\sigma^2(A_1, A_1)} = \sigma(A_1, A_1)$. With this new notation, we'll look at Lande's conceptualization of allometry. Note that since the allometry coefficient is a slope, and a slope is defined as a change in y over a change in x , we can write $\frac{\Delta \bar{z}_2}{\Delta \bar{z}_1}$. With this, we can consider a two-trait version of equation 16 (see equations 6 and 7) and write out the direct response of selection on body size as $\Delta \bar{z}_1 = \sigma^2(A_1, A_1) \beta_1$; note that this expression is simply the additive genetic variance for body size ($\sigma^2(A_1, A_1)$) multiplied by the strength of directional selection (β_1) acting on body size. Similarly, the correlated response of brainmass to selection on body size is $\Delta \bar{z}_2 = \sigma(A_1, A_2) \beta_1$; this expression is the additive genetic covariance between body size and brain size and the strength of directional selection acting on body size. Writing out these terms with respect to the way they scale allometrically ($\frac{\Delta \bar{z}_2}{\Delta \bar{z}_1}$) and doing some algebra yields,

$$\frac{\Delta \bar{z}_2}{\Delta \bar{z}_1} = \frac{\sigma(A_1, A_2) \beta_1}{\sigma^2(A_1, A_1) \beta_1} = \frac{\sigma(A_1, A_2)}{\sigma(A_1, A_1) \sigma(A_2, A_2)} \frac{\sqrt{\sigma^2(A_2, A_2)}}{\sqrt{\sigma^2(A_1, A_1)}} \quad (28)$$

The left term on the right side of equation 28 is a covariance divided by the standard deviation of traits 1 and 2. This is the definition of a correlation coefficient, often denoted as ρ . In this case, ρ is the additive genetic correlation, ρ_G , hence equation 28 can be written as

$$\frac{\Delta \bar{z}_2}{\Delta \bar{z}_1} = \rho_G \frac{\sigma(A_2, A_2)}{\sigma(A_1, A_1)} \quad (29)$$

Lande used experimental results from mice to determine the values of ρ_G and $\sigma(A_2, A_2)/\sigma(A_1, A_1)$; respectively these values are 0.68 and 0.524. Plugging these values into equation 29 gives us a value of

$$\frac{\Delta \bar{z}_2}{\Delta \bar{z}_1} = \rho_G \frac{\sigma(A_2, A_2)}{\sigma(A_1, A_1)} = 0.68 \times 0.524 = 0.356 \quad (30)$$

Recall Lande's original goal was to find a biological reason for why brain-body mass scaled with a ratio of 0.2–0.4 ($\frac{\Delta \bar{z}_2}{\Delta \bar{z}_1} = 0.2-0.4$). The calculated value in equation 30 falls within this range. Thus from equation 30, Lande concluded that the allometry of brain-body size among closely related species could be explained as a consequence of direct selection for body mass with brain size increasing as a result indirect selection. Allometry is widely studied in primates (e.g., [65] [43]); as genetic estimates become available, it will be interesting to see if other allometric patterns could be explained with respect to direct and indirect selection pressures.

6.6 Ecological morphology

In 1983, Arnold published an influential paper that connected selection gradients to ecological morphology ([5] also see [53]). Arnold’s main point was that equation 16 allowed researchers to directly measure the adaptive significance of a particular anatomical trait. His approach partitioned selection into two distinct components. One component of selection was the “performance gradient”, measured as the trait’s influence on some measure of performance (e.g., how lower limb length influences walking speed). The second component of selection was the “fitness gradient”, or how the measure of performance influences reproductive output or some other measure of fitness (e.g., how walking speed influences survival). This partition formalized the basic goal of ecological morphology: to understand morphological traits in the context of ecological selection pressures and evolutionary change. If we designate traits z , as above, performance measures as p , and fitness as F , then we relate the two gradients as depicted in Figure 4. Both gradients are estimable using partial regression (which controls for covariances between traits). Thus, the framework outlined by Arnold drew from the same statistical machinery as that of multivariate selection theory.

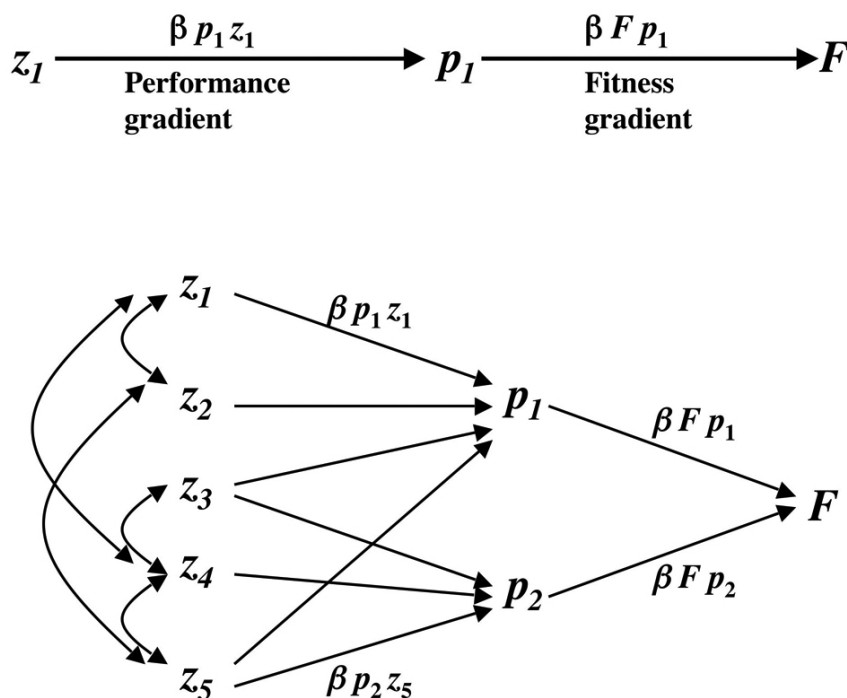


Figure 4: Arnold’s conceptualization of the relationship between morphology, performance, and fitness. The top part of the figure shows the basic pathway between a trait and fitness. The bottom shows the relationship between many traits, two measures of performance, and fitness. The performance gradients ($\beta p_i z_i$) and fitness gradients ($\beta F p_i$) can be measured using partial regression. The double-headed arrows represent phenotypic covariance.

Arnold’s paper is very important for formalizing and drawing connections between two fields that, at the time, were somewhat distinct: functional morphology and theoretical evolu-

tionary biology. This is not to say that practitioners of one field were ignorant of the basic tenets of the other field. Only that functional morphologists often made little mention of selection coefficients and variation in fitness, while those studying the theory of selection dynamics usually did not consider the nuances of form-function relationships. Arnold was able to specify a research program that connected how traits perform in a given ecological task with how such performance influences fitness. Arnold’s conceptualization grounded the raw material of evolutionary change—variation in traits—in a field that has *some* of its roots in typology and structuralism [4]. In addition, Arnold was able to provide an explicit methodology that allowed connections between proximate (e.g., ecological function) and ultimate (e.g., the adaptive process) domains in biology. While some of these ideas appear obvious in hindsight, Arnold provided an explicit methodology for merging empirical and theoretical research in the context of morphological evolution. Walker discusses an extension of Arnold’s approach [94].

6.7 Social behavior

While Arnold’s paper sought to connect morphological variation to fitness, a recent formalization has the promise to connect behavioral variation to variation in fitness [69] [97] [98]. As in other models of the phenotype, the approach of Moore et al., [69] breaks down a phenotypic trait (z_i) into two components: an additive genetic affect (a_i) and a general environmental effect ($e_{g,i}$). Thus we can write $z_i = a_i + e_{g,i}$. In this case we will view z_i as a behavior rather than an anatomical trait. In social behaviors, often one behavior provides the “environment” that influences another behavior. For example, a dominance threat displayed by one individual is a behavior that often elicits a behavioral response (e.g., submissive gestures) in a different individual. In this context, one individual’s behavior is an environmental influence on another individual’s behavior. To this end, we can write out how another behavior (z_j) influences our behavioral trait z_i as

$$z_i = a_i + e_{g,i} + \psi_{ij}z_j \tag{31}$$

In this case, the new terms are the other individual’s behavior (z_j) and the effect (ψ_{ij}) that this behavior has on our focal trait z_i . This is illustrated in Figure 5.

With this basic framework, let us consider the case of dominance and submissive behaviors. In many such cases, agonistic threats influence the subordinate’s level of submissive response and, in turn, the subordinate’s response mediates the intensity of the aggressor’s threats. In this case there is a reciprocal effect of one trait’s expression on another trait. Putting this situation into mathematical notation, let’s denote dominance threats as z_2 and submissive displays as z_1 . Both of these traits are capable of being displayed by all individuals but they are treated as separate characters. We will write out expressions for how dominance displays are composed of direct genetic and environmental effects and an indirect effect. They are (the ‘ indicates the trait of another individual),

$$z_1 = a_1 + e_{g,1} + \psi_{12}z_2' \tag{32}$$

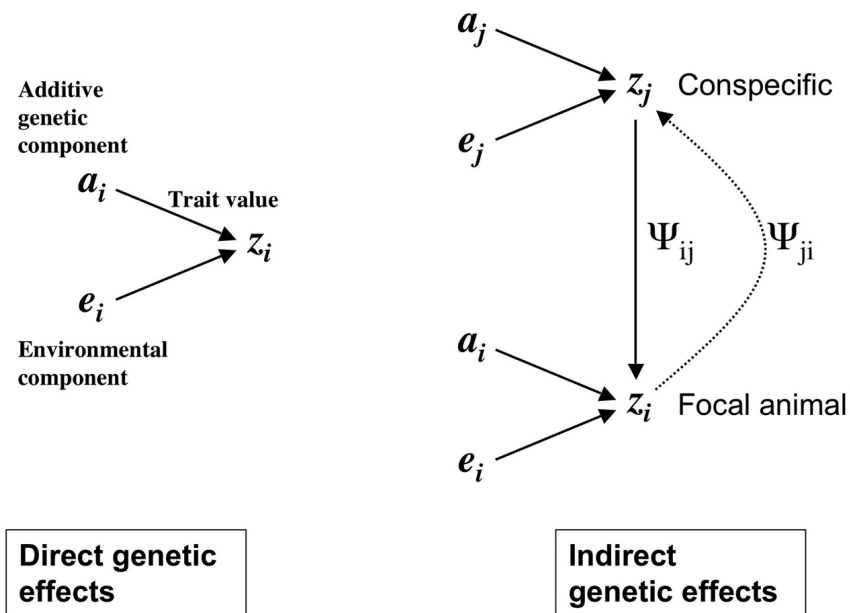


Figure 5: The logic behind social selection models. If there is only a single individual, a trait z_i is composed of additive genetic component (a_i) and an environmental component (e_i). When there are social effects, the trait z_i is similarly influenced by an additive and environmental effect, but in addition, a conspecific's trait (z_j) (or behavior) also influences the expression of the focal individuals trait (z_i). The term Ψ_{ij} measures the magnitude of the effect that z_j has on z_i . When z_i also has an effect on z_j (dotted line) we would note the effect of z_i on z_j as Ψ_{ji} .

$$z_2 = a_2 + e_{g,2} + \psi_{21}z_1' \quad (33)$$

From equation 32, we see that submissive displays are influenced by additive genetic effects (a_1), general environmental effects ($e_{g,1}$), and the *strength* of the influence, denoted as ψ_{12} , that dominance displays (z_2) have on submissive displays. Equation 33 gives the components of dominance behaviors in light of the expression of submissive behaviors. Moore et al., [69] derived a multivariate expression for how average trait values change under selection due to *indirect genetic effects*, such as the social environment provided by another individual. This expression is

$$\Delta\bar{z} = \mathbf{C}_{\alpha p} \beta (\mathbf{I} - \Psi)^{-1} \quad (34)$$

In this equation, the mean change in an array of behaviors is determined by direct selection on each trait (the β term) as well as by a matrix which specifies the strength of the effect of outside behaviors—the indirect genetic effects; this is denoted by $(\mathbf{I} - \Psi)^{-1}$. The term \mathbf{I} denotes an identity matrix, which is a matrix of ones along the diagonal and zeros on the off-diagonal and Ψ is a matrix of coefficients that specify the strength of the indirect effect. We have replaced the familiar \mathbf{G} matrix with $\mathbf{C}_{\alpha p}$ —a matrix of the correlations between breeding value and corresponding phenotype. In most situations, we can write the correlation between breeding value and phenotype ($\rho_{\alpha,P}$) in terms of V_A (that is, $V_A = \rho_{\alpha,P}$). However, when a trait is influenced by indirect effects, then we can no longer use V_A as a stand-in for $\rho_{\alpha,P}$. This is because our trait in question, z , is now a function of its own direct additive genetic and environmental effects, plus the additive genetic and environmental effects of its conspecific (see Figure 4). The indirect genetic effects enhance or diminish the correlation between the breeding value and phenotype of the focal individual. In this case, one needs to substitute the actual values for $\rho_{\alpha,P}$ into $\mathbf{C}_{\alpha p}$ in lieu of using V_A as we do in the \mathbf{G} matrix.

It is helpful to write out the the average change in a single trait. From equation 34, we can write

$$\Delta\bar{z}_1 = [G_{ff} \beta_1 + G_{fc} \beta_2] + \psi_{12}[G_{fc} \beta_1 + G_{cc} \beta_2] \quad (35)$$

From equation 35, we see that the average change in trait z_1 is specified by direct selection on the trait and its covariance with another trait (the bracketed terms on the left), plus the indirect genetic effects (the bracketed terms on the right). The new terms are G_{ff} , the additive genetic variance in the focal animal’s trait; G_{cc} , the additive genetic variance in the conspecific’s trait; and, G_{fc} , the genetic covariance between direct additive effects and indirect additive effects. If the term G_{fc} is positive, alleles that cause an increase in the focal animal’s trait also increase the conspecific’s trait. The key to understanding this equation is to note that a social behavior is both an environmental effect and a genetic effect. That is, the behavior creates an environment that influences a particular trait, and the behavior itself usually has some genetic basis (i.e., it is partly heritable). This model of indirect genetic effects is a generalization of a kin selection model. It is possible to replace ψ with “ r ”, the coefficient of relatedness; when we do this, we have a model of kin selection that incorporates not only the influence of a relative’s behavior on the focal animal’s trait, but also the influence of the focal animal’s own additive genetic effects on its own trait. Moreover, this model does not ignore heritability of traits—a key parameter that is not included in most conceptualizations of Hamilton’s rule (see [28]). Thus, equation 34 provides a very

general framework for understanding the influence of the allele's carried by conspecifics on the expression of a trait in a focal animal.

Finally a topic of recent interest that we do not cover here is phenotypic integration (often known as morphological integration). Perhaps more so than any other topic covered here, phenotypic integration has a long-standing tradition in primate biology and some of the first studies of phenotypic integration were conducted on owl monkey teeth [70]. phenotypic integration draws from the field of quantitative genetics since it concerns several (presumably polygenic) traits and how these traits evolve to become interdependent. In particular, phenotypic integration manifests as strong phenotypic covariation due to traits which are functionally or developmentally interdependent. As such, phenotypic integration involves an investigation of the \mathbf{G} and \mathbf{P} matrices. Cheverud [27], among others, showed that there is a strong correlation between \mathbf{G} and \mathbf{P} . Due to this, many studies of phenotypic integration only examine patterns of phenotypic covariation (e.g., [1] [63] [85]). Ackermann and Cheverud [2] provide an excellent review of phenotypic integration studies in primates, and methodological and theoretical approaches are covered in [33] [24] [68]. The general framework of phenotypic integration has recently been extended to behavioral traits in an approach commonly known as “behavioral syndromes.” This approach seeks correlations in particular behaviors among individuals in a population; the framework for studying suites of behaviors is accommodated using a phenotypic (or genetic) correlation matrix (reviewed in [87]).

7 Quantitative Genetic Studies of Primates: Some Examples

Some early and influential studies of quantitative genetics were first conducted on primate species, notably in the work of Cheverud (e.g., [30] [31]). Since then, the estimation of quantitative genetic parameters on wild and free-ranging primates has been sparse, but growing. In this section we review studies that have estimated one or more of the various parameters, or associated parameters, in equation 16 in wild or free-ranging primate populations. These include selection differentials, selection gradients, heritability, and genetic correlations. Our review is not exhaustive; instead, we focus on recent studies that illustrate the diverse approaches to understanding adaptive evolution in non-human primates. There is a large literature on heritability/quantitative genetics for captive primates as well as humans; we don't review these topics.

7.1 Measuring selection in primate populations

Although methods for measuring selection have been around for a long time, there have been few applications to wild primate populations. We summarize some of the most recent

studies and the main results of these studies are provided in Table 3. De Gusta et al., [39] figured out a way to measure viability selection in a population of wild Howler monkeys (*Alouatta palliata*) from Barro Colorado island. They were able to determine a proxy of a selection differential by measuring the bucco-lingual width of the first upper molar (UM1). The measurements came from a collection of crania “found dead” of natural causes on the Island. The crania were sexed and aged based on standard techniques. Each cranium was assigned to one of five dental stages based on the eruption sequence of other teeth in the crania. Their choice of UM1 width was based on the fact that this trait has been shown to be partly heritable and that the molar crown is fully formed soon after birth and therefore does not exhibit growth related changes. Since all animals they looked at were dead, and eruption patterns provide an estimate of relative age, then these data provided estimates of age at death—a measure of fitness. Their phenotypic measurement doesn’t change with age but shows natural variation, some of which may correlate with survival. De Gusta and colleagues looked at the width of UM1 against different age classes and showed that animals with smaller UM1 widths were more often found dead than animals with larger UM1 widths. They interpreted this as viability selection against animals with smaller UM1s. Specifically, they found that the highest mortality, and hence strongest selection, was against animals in the 6-12 month stage, a period that corresponds to weaning. As they note, they were only able to measure this particular trait’s correlation with fitness and were not able to parse selection into direct and indirect components, nor account for variation in maternal health. Nevertheless, De Gusta et al.’s study is interesting because it demonstrates that selection can be estimated in non-obvious ways—a collection of dead animals. This approach is one of the first to measure phenotypic selection in a wild population of primates.

Another recent study of viability selection used demographic data to estimate selection on foot length in a wild population of Verreaux’s sifaka (*Propithecus verreauxi verreauxi*). One of the goals of this study was to examine if foot length, which scales with negative allometry to body size during growth, was associated with survival [62]. In this regard, Lawler [62] sought to test if foot size was adaptive. Since foot size is a trait that changes with age it is necessary to remove growth related variation. This was accomplished by regressing foot size against age. The residuals from this regression provide a measure of age-corrected size variation. Other age-corrected traits included humerus length, radius length, hand length, femur length, and tibia length. Multiple regression was used in order to estimate the magnitude of direct selection (i.e., the selection gradient, β) acting on each trait. A discrete measure of fitness was used as the dependent variable: alive past age of eight or dead before age of eight. All traits were standardized to have a mean of 0 and a standard deviation of 1 and fitness was scaled to produce relative fitness. The resulting directional selection coefficients indicated that strong selection was acting on foot length ($\beta = 0.119$). Standardizing the traits allows us to interpret the selection coefficient as follows: relative fitness will increase by 11.9% for every increase in one standard deviation in foot length. From these findings, Lawler proposed that increased foot length (relative to body size) plays a role in allowing younger animals to safely leap between substrates; younger animals with relatively shorter feet have lower survivorship.

A similar study on this same population estimated the strength and type of intrasexual

selection in male Verreaux’s sifaka [61]. In this study, the goal was to determine which traits were important determinants of male fertility. As such, the measure of fitness was the relative rate of offspring sired by males over a given sample period (that is, the number of offspring produced produced by each male divided by the number of years he spent in the sample period). The traits in the analysis were body mass, canine length, arm shape, torso shape, and leg shape (the latter three traits were principal components of numerous linear measurements). Lawler et al., [61] estimated different types of selection acting on these five traits with respect to male fertility. They found that significant directional selection acts on leg shape and stabilizing selection (γ_{ii}) acts on body mass. In addition, there was negative correlational selection (γ_{ij}) for leg shape and body mass. These patterns of selection were interpreted with respect to male mating competition. Male mating contests take place in the trees and males with larger values of leg shape (which means longer legs and larger thigh circumference) and “stream-lined” body mass were likely better equipped to engage in such contests. Thus, important determinants of male fertility in this population corresponded to traits that enhanced locomotor contests during the mating season.

So far, we have been considering the effects on selection on anatomical traits. However, as discussed above, selection can also target life history traits—traits like age at sexual maturity, survivorship in age-class k , etc. Alberts and Altmann [3] constructed an age-based matrix population model using long-term data from female savana baboons (*Papio papio*). The life cycle of the baboon was similar to that of Figure 2, but contained 27 stages, of which the first 26 were ages 1 through 26, and the 27th stage contained a “self-loop” corresponding to a stage of animals aged 26 years and above. The resulting projection matrix was a 27×27 matrix. It is possible to calculate the directional selection acting on age-specific survival and growth by calculating the “sensitivity” of fitness to transitions (i.e., surviving the year and growing onto the next age) between particular age classes. Technically, Alberts and Altmann [3] calculated elasticities, which are scaled sensitivities, but we present sensitivity values here. Using the right hand side of equation 27, sensitivity values for all age class transitions were calculated. The overall pattern that emerges from these calculations is that strong directional selection acts on the youngest age classes. For example, the directional selection coefficient ($\frac{\partial \lambda}{\partial a_{21}}$) acting on the transition from age 1 to age 2 (this would correspond to the G_1 transition in Figure 3) is 0.127. Coefficients for other transitions between younger age classes are similarly valued around 0.11. The strength of directional selection drops off quite steeply for transitions between older age classes, a pattern seen in most other long-lived vertebrates [22].

7.2 Measuring heritability and genetic correlations in primate populations

In this section, we examine studies that estimated heritability and genetic correlations in wild or free-ranging primates. Because heritability is proportional to additive genetic variance, it is a key parameter governing the evolutionary response to selection (see equations 1 and

Table 3: Summary of studies on wild/free-ranging primate populations that estimated selection

Trait(s) involved	Parameter	Reference	Major finding
Upper molar width	S	[39]	Animals with wider molars have higher survivorship
Foot length	β	[62]	For animals 1 to 8yrs, longer feet is associated with higher survival
Leg shape	β	[61]	Male fertility iss associated with longer, stronger legs
Body mass	γ_{ii}	[61]	Male fertility is associated with intermediate body mass
Body mass and leg shape	γ_{ij}	[61]	Correlational selection acts on body mass/leg shape
Survival in age class k	$\frac{\partial \lambda}{\partial a_{ij}}$	[3]	Directional selection is strongest at younger age classes

15). Accordingly, several researchers have sought to estimate the heritability and genetic correlations of morphological and life history traits in wild primate populations. These estimates provide insight into the standing levels of additive genetic variation and covariation for particular traits, thus helping to determine how a population will adapt to selection pressures. Table 4 provides some examples of studies that have estimated heritability and genetic correlations in wild and free-ranging primates.

One of the first large-scale studies of morphological heritability on a wild primate population was conducted by Cheverud and Dittus [32]. This population of Toque macaques (*Macaca sinica*) has been studied for over 30 years at Pollonuaruwa in Sri Lanka. Cheverud and Dittus sought to estimate the heritability of 27 morphometric measurements, which included traits such as crown-rump length, arm length, leg length, tail length, and head length. The long-term behavioral observations made it possible to determine 172 maternal-offspring pairs from 39 different maternal lineages. Because the data come from animals of different ages as well as from sons and daughters, the authors first removed any age and sex-effects via two steps; first, they fit a flexible spline regression to each trait by age (and sex) in order to identify natural "breaks" in the trait distribution with respect to age. The spline regression suggested three age classes, and thus the subsequent step was to perform linear regression of each trait against that particular age class (and sex). This resulted in sets of residual values that were then used to estimate heritabilities of the traits. To estimate heritabilities they linearly regressed the offspring's trait value on the corresponding mother's trait value using linear regression, and they also used "pedigree methods;" this latter method takes advantage of the fact that trait values covary among all relatives, not just mother-offspring pairs [56]. Using linear regression, the heritability is 2x the regression slope value, since mother's only

contribute 1/2 of the additive genetic variance to offspring trait values. They found that 23 of the 27 heritability estimates were significantly different from zero using the regression method and all heritability estimates were significant using pedigree methods. The average heritability among all traits was 0.51 using mother-offspring regression and 0.56 using the full pedigree. These results suggest that a significant amount of anatomical traits, particularly lengths of long bones, spine, and head, and the circumferences of limb elements contain a significant amount of variation that is inherited. As such, this population has the capacity to respond to selection on numerous skeletal elements or overall body form itself [32].

Another study that looked at heritability in a wild primate population was that of Lawler [62]. This study is outlined above with respect to estimating selection pressures that act on limb elements on Verreaux’s sifaka. In addition to estimating selection gradients, Lawler also estimated the heritability of limb elements and hand/foot length. He used 146 father-offspring pairs. The use of father-offspring pairs can reduce the influence of maternal effects (see section 2), whereby some of the phenotypic variation in traits is due to differences among groups of siblings born to different mothers. Lawler found that heritability values were lowest in the hands and feet (0.116 for the foot and 0.21 for the hand) and the values increased in magnitude moving up the limb from hand/foot to lower arm/leg to upper arm/thigh. Lawler interpreted this pattern as partially reflecting the action of past and current selection pressures acting on hands and feet (recall from above that positive directional selection acts most strongly on foot length) based on the idea that selection should erode the amount of additive genetic variance and hence lower heritability as it acts to change the mean value of a trait across generations. However, the heritability values for these traits have very large standard errors so this study would benefit from a more robust analysis of heritability.

The above two studies focused on skeletal traits but a more recent study looks at the heritability of life history traits in free-ranging macaques. In a recent study of the quantitative genetics of Cayo Santiago macaques (*Macaca mullata*), Blomquist [10] examined the heritability of three traits that are strongly associated with fitness: lifetime reproductive success (LRS), lifespan, and individual rate of increase (λ_i). LRS is the total number of offspring born to females regardless of the offspring sex or whether the offspring died early into its life; lifespan is measured in years and calculated from long-term demographic observations. λ_i is an individual measure of the growth rate—not “growth” in a skeletal sense, but growth in a demographic sense— λ_i measures the rate and timing of offspring production for individuals as calculated by an individual projection matrix (in section 6.3 we discuss the population projection matrix, which contains the average probabilities of survival, growth and reproduction for a population from which the population growth rate, λ , can be calculated; it is possible to construct analogous projection matrices for individuals and then use these matrices to calculate the individual growth rate, λ_i). The difference between LRS and λ_i is that the latter takes into account not only total number of offspring produced but *when* in the life cycle these offspring were born (see discussion on Fitness in section 4). Blomquist estimated these “fitness traits” using two sets of females—those that lived “whole” lives from birth to a natural death who may or may not have reproduced (the uncensored group), and a nested set of the uncensored group of females which comprised females who lived “whole” lives from birth to natural death but reproduced at least once in their lives (the censored

group). Blomquist estimated the life history traits by writing the phenotypic variation (V_P) as the sum of other components of variation that included birth cohort (V_C), maternal effects (V_M), social group (V_G), additive genetic (V_A), and residual variation (V_R) (see section 2). When values for each variance component were estimated, it was possible to calculate heritability by taking the ratio of additive genetic variance to total phenotypic variance ($h^2 = V_A/V_P$). Heritability of LRS, lifespan, and λ_i was 0.38, 0.36, and 0.43 respectively using the uncensored group. When the censored group was used the heritability estimates were roughly half the value of the uncensored estimates. The difference between these sets of estimates can be attributed to the fact that a large portion of genetic variation is due both survival and recruitment; that is, when females who *have* reproduced are the only ones included in the analysis, this censored data set purposely omits a lot of females (and hence a lot of variation) who didn't survive to reproductive age. Hence, the total amount of phenotypic variation and additive genetic variation is reduced by not including females who died prior to reproduction. The environment at Cayo Santiago is relatively homogenous and free from predators and factors causing nutritional stress. In this case, if there is lots of additive genetic variation for survival in early years, a homogenous environment will create conditions in which this additive genetic variance for survival gets exposed and animals with "bad" genotypes die off (since there few environmental factors that cause young animals to die due to predation and/or nutrition—that is, there is reduced environmental variation). This results in moderately high heritabilities for uncensored fitness traits because of lot of the variation observed in the population is due to genetic differences with respect to survival and reproduction, not environmental differences.

The above studies focused on estimating heritability for various morphometric and life history traits. Additional studies of free-ranging and captive primates have estimated genetic (ρ_G) and phenotypic (ρ_P) correlations among traits. We will discuss two such studies. Hlusko et al., [50] used quantitative genetic techniques to estimate heritability and genetic correlations among dental measurements and trunk length in a captive population of Hamadryas baboons (*Papio hamadryas*), which is comprised of different subspecies that interbreed (*P. h. anubis* and *P. h. cynocephalus*). This population resides at the Southwest Foundation for Biomedical Research. It has been studied for over 20 years and mating designs are implemented in this population in order to ensure non-inbred animals. Hlusko et al., used a model of phenotypic variation in which phenotypic variance was written as the sum of additive genetic variation and environmental variation ($V_P = V_A + V_E$) but tested for the effect of different covariates (e.g., sex, age, percent of subspecies admixture, etc.) that may also influence phenotypic variation. They tested whether genetic correlations exist between trunk length and several measurements of the upper and lower second molar. It was possible to do this by considering a two-trait version of their phenotypic decomposition into additive and environmental components. In this case, the output consists of a phenotypic variance/covariance matrix (\mathbf{P}) which can be broken down into additive and environmental components. They found significant genetic correlations between upper molar area and trunk length ($\rho_G = 0.44$) as well as lower molar area and trunk length ($\rho_G = 0.56$). Phenotypic correlations between these traits were much lower. Their results suggest that about 19–42% of the additive effects are shared between trunk length and the area of the upper and lower molar. Additional genetic correlations were found to exist between linear measurements of crown area and trunk

length with the majority of the significant correlations existing between buccolingual width rather than mesiodistal length. As Hlusko et al., (page 424) note, the mechanisms that link up these two traits are likely due to, “latent rather than specific genetic factors.” The low phenotypic correlation, in spite of the modest genetic correlation, is likely due to the various differences in the way these two traits grow and/or are influenced by nongenetic factors. Hlusko et al., note the relevance of this finding in that molar length is often observed to increase in mammalian lineages whereas molar width does not. Their results suggest that a genetic correlation between body size and molar width constrains molar dimension to only increase in length not width. Further, they note that in phyletically dwarfed lineages tooth width decreases much more rapidly than tooth length—this is possibly due to selection for smaller body size and a correlated response in tooth width.

Genetic correlations have also been estimated for life history traits. Using the same large-scale dataset for Cayo Santiago Macaques, Blomquist [11] estimated the genetic correlation between age at first reproduction (AFR) and adult survival in females. Both AFR and lifespan are heritable in this population [10] [12]. Motivating this study is the widespread theoretical and empirical pattern of trade-offs among life history traits. Common trade-offs in life history theory include current versus future reproduction and a trade off between survival and reproduction. As Blomquist discusses, it is particularly important to estimate the genetic correlations among life history traits because phenotypic correlations among such traits often do not provide any indication of their underlying genetic covariation; this is usually due to the confounding effect of environmental correlations and environmental variation that influences life history traits. Because primates are long-lived creatures, a lot of total adult fitness depends on surviving from year to year. In the macaque population at Cayo Santiago, there is a very strong and positive correlation between lifespan and lifetime reproductive output. Due to this positive correlation Blomquist sought to determine if there was a genetic correlation between AFR and survival. In this case, a *positive* genetic correlation would be evidence for a tradeoff since a later age at first reproduction would ostensibly enhance survival, whereas reproducing earlier would diminish survival—the trade-off. He defined four ages that could be used as survival milestones: 11, 16, 21, 26. Thus Blomquist was looking for a positive genetic correlation between AFR and survival rates to these four ages. Blomquist estimated the genetic correlation between AFR and adult survival using a multivariate model in which phenotypic variation is decomposed into additive genetic and residual variance ($V_P = V_A + V_R$). He obtained the additive genetic and residual covariances between traits and converted these covariances into correlations. The phenotypic correlations between AFR and the four survival milestones were relatively small ($\rho_P = 0.109$, $\rho_P = 0.132$, $\rho_P = 0.101$, $\rho_P = 0.083$, respectively) and only one correlation was significant at the 0.05 level. However, strong and positive additive genetic correlations existed between AFR and the four survival milestones: $\rho_G = 0.590$, $\rho_G = 0.595$, $\rho_G = 0.476$, $\rho_G = 0.706$. The additive genetic correlation between AFR and age 11 was not significant at the 0.05 level, but the other three correlations were significant. These results suggest a genetically mediated trade-off between reproductive maturation and survival. Females who reproduce earlier are less likely to survive to later ages. In fact, Blomquist was able to numerically calculate the magnitude of the trade-off: the cost of reproducing 1 year earlier in life results in a loss of approximately 11 months of adult life. This is the first empirical study of nonhuman

primates that documents a key principle of life history theory, namely that genetic mediated trade-offs exist between different components of fitness.

Table 4: Summary of studies on wild/free-ranging primate populations that estimated heritability and genetic correlations

Trait(s) involved	Parameter	Reference	Value of parameter/comments
Various traits	h^2	[32]	average h^2 was 0.53, lower values for distal elements
Foot length	h^2	[62]	heritability was 0.16 and higher for other limb elements
LRS, λ_i , lifespan	h^2	[10]	$h^2 = 0.38, 0.36, 0.43$ respectively (uncensored group)
Molar area and trunk length	ρ_G	[50]	Positive genetic correlation among these two traits
Age at first reproduction and survival	ρ_G	[11]	Positive genetic correlation among these life history traits

8 Conclusions

Because the basic properties of adaptive evolution are few: heritable variation \times selection pressure = change in phenotype, it is possible to modify equation 16 to capture almost any specific evolutionary scenario that involves phenotypic change. Our section showing the extensions of equation 16 to specific evolutionary scenarios captures only a portion of the diverse approaches to studying phenotypic evolution in a quantitative genetic framework. In contrast, our section reviewing quantitative genetic studies wild/free-ranging primate populations are few; the lack of application is likely due to estimation problems: gathering sufficient data from which to estimate heritability and/or selection pressures can be daunting or unfeasible. In this largely theoretical overview, we have omitted a section on how to actually estimate the quantitative genetic parameters that we discuss. The topic of estimation is a very mathematically dense literature. However, several good reviews are found in [64] [56] [96].

It's important to remember three rules of thumb when applying any of these multivariate procedures to actual data: 1) *Be a good biologist*. By this, we mean don't be wooed by the multivariate analysis; you still need to interpret your data. It is especially imperative to determine the causal structure that creates a non-zero relationship between fitness and phenotype. The statistical relationship is vacuous without biological input from many careful hours of observation and thought; 2) *Remember what you left out*. Your analysis is only

as good as the traits you’ve included in the model. It is always possible that some unmeasured trait influences the traits that you have actually measured—you need to be mindful of this possibility; and, 3) *Do not extrapolate wildly*. Equation 16 provides information on evolutionary trajectories over the short-term and these trajectories are specific to the population you’re studying. Selection pressures and additive genetic variances (and covariances) change due to a variety of circumstances. To suggest that trait heritabilities or trait-trait correlations from your population likely apply to other populations invites well-deserved scrutiny.

The above might seem like a lot of information to the non-initiate. However, like most topics in biology, it is only possible to scratch the surface of the field in a short review such as this. We have left out many important topics such as phenotypic plasticity, maternal effects, and how other evolutionary forces (e.g., non-random mating, drift, etc.) impinge on the theory above. This review should serve only as an introduction to the field, not as a one-stop source. The reader is highly encouraged to consult [88] [82] [83] [64] [37] [80] for further information.

9 Acknowledgements

We thank the following institutions for supporting our research: Yale University, University of Illinois, Boston University, University of Missouri, James Madison University, and the National Science Foundation (BCS 0820298, DBI 0305074). RRL thanks the mathematical ecology group at Woods Hole Oceanographic Institution, in particular Hal Caswell, Mike Neubert, Christine Hunter, Tin Klanjscek, and Petra Klepac. We also thank Hal Caswell for giving us feedback on calculating sensitivities. RRL heard the quote “How much of the real world can we throw away and still get the right answer” in an Ecology class taught by Leo Buss. We thank one anonymous reviewer for their helpful comments that greatly improved the clarity of this manuscript.

References

- [1] Ackermann RG. Patterns of covariation in the hominoid craniofacial skeleton: Implications for paleoanthropological models. *J Hum Evol* 2002; 43: 167-187.
- [2] Ackermann RR, Cheverud JM. Morphological integration in primate evolution. In: M Pigliucci, K Preston eds. Oxford, United Kingdom 2004 pp. 302-320.
- [3] Alberts SC, Altmann J. Matrix models for primate life history analysis. In: PM Kappeler, M Pereira eds. Chicago, USA 2004. pp. 66-102.
- [4] Amundson R. The changing role of the embryo in evolutionary thought: Roots of Evo-Devo. Cambridge United Kingdom: Cambridge University Press 2007.
- [5] Arnold SJ. Morphology, performance, and fitness. *Am Zool* 1983; 23: 347-361.
- [6] Arnold SJ. Constraints on phenotypic evolution. *Am Nat* 1992; 140: S85-107.
- [7] Arnold SJ. . Multivariate inheritance and evolution: A review of concepts. In: CRB Boake, ed. *Quantitative Genetic Studies of Behavioral Evolution*. Chicago, USA 1994; pp. 17-48.
- [8] Arnold SJ, Wade MJ. On the measurement of natural and sexual selection: theory. *Evolution* 1984; 38: 709-719.
- [9] Benton TG, Grant A. Evolutionary fitness in ecology: Comparing measures of fitness in stochastic, density-dependent environments. *Evol Ecol Res* 2000; 2: 769-789.
- [10] Blomquist GE. Heritability of individual fitness in female macaques. *Evol Ecol* 2010; 24: 657-659.
- [11] Blomquist GE. Trade-off between age of first reproduction and survival in a female primate. *Biol Letters* 2009b; 5: 339-342.
- [12] Blomquist GE. Fitness-related patterns of genetic variation in rhesus macaques. *Genetica* 2009c; 135: 209-219.
- [13] Boake CRB. ed. *Quantitative Genetic Studies of Behavioral Evolution*. Chicago, USA. University of Chicago Press 1994.
- [14] Bock WJ. The definition and recognition of biological adaptation. *Am Zool* 1980; 20: 217-227.
- [15] Brodie ED, Moore AJ, Janzen FJ. Visualizing and quantifying natural selection. *Trends Ecol Evol* 1995; 10: 313-318.
- [16] Brodie ED, Janzen FJ. On the assignment of fitness values in statistical analyses of selection. *Evolution* 1996; 50: 437-442.
- [17] Brommer JE. The evolution of fitness in life history theory. *Biol Rev* 2000; 75: 377-404.

- [18] Brommer JE, Merila J, Kokko H. Reproductive timing and individual fitness. *Ecol Letters* 2002; 5: 802-810.
- [19] Caswell H. Life history theory and the equilibrium status of populations. *Am Nat* 1982; 120: 317-339.
- [20] Caswell H. Theory and models in ecology: A different perspective. *Ecol Model* 1988; 43: 33-44.
- [21] Caswell H. *Matrix Population Models*. Sunderland, USA; Sinauer Press 1989.
- [22] Caswell H. *Matrix Population Models*. Sunderland, USA; Sinauer Press 2001.
- [23] Charlesworth B. *Evolution in Age Structured Populations*. Chicago, USA; University of Chicago Press 1994.
- [24] Chernoff B, Magwene PM. *Morphological Integration: Forty years later*. Morphological Integration, Chicago USA, University of Chicago Press 1999.
- [25] Cheverud JM. Relationships among ontogenetic, static, and evolutionary allometry. *Am J Phys Anthropol* 1982; 59: 139-149
- [26] Cheverud JM. Quantitative genetics and developmental constraints on evolution by selection. *J Theor Biol* 1984; 110: 155-171.
- [27] Cheverud JM. A comparison of genetic and phenotypic correlations. *Evolution* 1988; 42: 958-968.
- [28] Cheverud JM. Evolution in a genetically heritable social environment. *PNAS* 2003; 100: 4357-4359.
- [29] Cheverud JM. Genetic architecture of quantitative variation. In CW Fox, JB Wolf, eds. *Evolutionary Genetics: Concepts and Case Studies*. Oxford, United Kingdom. pp. 288-309.
- [30] Cheverud JM, Buikstra JE. Quantitative genetics of skeletal nonmetric traits in the rhesus macaques on Cayo Santiago. II Phenotypic, genetic, and environmental correlations between traits. *Am J Phys Anthropol* 1981; 54: 51-58.
- [31] Cheverud JM, Buikstra JE. Quantitative genetics of skeletal nonmetric traits in the rhesus macaques on Cayo Santiago. III Relative heriability of skeletal nonmetric and metric traits. *Am J Phys Anthropol* 1982; 59: 151-155.
- [32] Cheverud JM, Dittus WPJ. . Primate population studies at Polonaruwa 2. Heritability of body measurements in a natural population of Toque macaques (*Macaca sinica*). *Am J Primatol* 1992; 27: 145-156.
- [33] Cheverud JM, Wagner GP, Dow MM. Methods for the comparative analysis of variation patterns. *Am Zool* 1989; 38: 201-213.

- [34] Coddington JA. Cladistic tests of adaptational hypotheses. *Cladistics* 1988; 4: 3-22.
- [35] Conner J. Field measurements of natural and sexual selection in the fungus beetle *Bolitotherus cornutus*. *Evolution* 1988; 42: 736-749.
- [36] Conner JK. How strong is natural selection? *Trends Ecol Evol* 2001; 16: 215-217.
- [37] Conner JK, Hartl DL. *A Primer of Ecological Genetics*. Sunderland, USA, Sinauer Press 2004.
- [38] Coulson T, Benton TG, Lundberg P, Dall SRX, Kendall BE, Gaillard JM. Estimating individuals contributions to population growth: evolutionary fitness in ecological time. *Proc Roy Soc Lond B* 2006; 273: 547-555.
- [39] DeGusta D, Everett MA, Milton K. Natural selection on molar size in a wild population of howler monkeys (*Alouatta palliata*). *Proc R Soc Lond B (Suppl.)* 2004; 270: S15-S17.
- [40] Fairbairn DJ. Allometry for sexual size dimorphism: Pattern and Process in the coevolution of body size in males and females. *Ann Rev Ecol Syst* 1997; 28: 659-687.
- [41] Falconer DS, MacKay TFC. *Introduction to Quantitative Genetics*. London, United Kingdom, Longman Publishing 1996.
- [42] Fisher DC. Evolutionary morphology: Beyond the analogous, the anecdotal, and the ad Hoc. *Paleobiol* 1985; 11: 120-138.
- [43] Fleagle JG. *Primate Evolution and Adaptation*. New York, USA, Academic Press 1998.
- [44] Fuller RC, Houle D, Travis J. Sensory bias as an explanation for the evolution of mate preferences. *Am Nat* 2005; 166: 437-466.
- [45] Futuyma DJ. *Evolutionary Biology*. Sunderland, USA, Sinauer Press 1998.
- [46] Gillespie JH. *Population Genetics: A Concise Guide*. Baltimore, USA, Johns Hopkins Press 2004.
- [47] Gordon AD. Scaling of size and dimorphism in primates: Microevolution. *Int J Primatol* 2006; 27: 27-61.
- [48] Gould SJ, Vrba ES. Exaptation: A missing term in the science of form. *Paleobiol* 1982; 8: 4-15.
- [49] Harvey PH, Pagel MD. *The Comparative Method in Evolutionary Biology*. Oxford, United Kingdom, Oxford University Press 1991.
- [50] Hlusko LJ, Lease LR, Mahaney MC. Evolution of genetically correlated traits: Tooth size and Body size in Baboons. *Am J Phys Anthropol* 2006; 131: 420-427.
- [51] Jones JH demogR: A package for the construction and analysis of age-structured demographic models in R. *J Stat Software* 2007; 22(10): 1-28.

- [52] Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P, Beerli P. The strength of phenotypic selection in natural populations. *Am Nat* 2001; 157: 245-261.
- [53] Kingsolver JG, Huey RB. Introduction: The evolution of morphology, performance, and fitness. *Integr Comp Biol* 2003; 43: 361-366.
- [54] Kirkpatrick M. Evolution of female choice and male parental investment in polygynous species: The demise of the “sexy son.” *Am Nat* 1985; 125: 788-810.
- [55] Kokko H, Jennions MD, Brooks R. Unifying and testing models of sexual selection. *Ann Rev Ecol Evol Syst* 2006; 37: 43-66.
- [56] Kruuk LEB. Estimating genetic parameters in natural populations using the ‘animal model.’ *Phil Trans R Soc Lond B* 2004; 359: 873-890.
- [57] Lande R. Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry. *Evolution* 1979; 33: 402-416.
- [58] Lande R. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 1980; 34: 292-305.
- [59] Lande R. A quantitative genetic theory of life history evolution. *Ecology* 1982a; 63: 607-615.
- [60] Lande R, Arnold SJ. The measurement of selection on correlated characters. *Evolution* 1983; 37: 1210-1226.
- [61] Lawler RR, Richard AF, Riley MA. Intrasexual selection in Verreauxs sifaka (*Propithecus verreauxi verreauxi*). *J Hum Evol* 2005; 48: 259-277.
- [62] Lawler RR. Sifaka positional behavior: Ontogenetic and quantitative genetic approaches. *Am J Phys Anthropol* 2006; 131: 261-271
- [63] Lawler RR. Morphological integration and natural selection in the postcranium of Verreaux’s sifaka. *Am J Phys Anthropol* 2008; 136: 204-213.
- [64] Lynch ML, Walsh JB. *Genetics and the Analysis of Quantitative Traits*. Sunderland, USA Sinauer Press, 1998.
- [65] Martin RD. *Primate Origins and Evolution*. Princeton, USA, Princeton University Press, 1990.
- [66] McGraw JB, Caswell H. Estimation of individual fitness from life history data. *Am Nat* 1996; 147: 47-64.
- [67] Mitchell-Olds T, Shaw RG. Regression analysis of natural selection: Statistical inference and biological interpretation. *Evolution* 1987; 41: 1149-1161.

- [68] Mitteroecker P, Bookstein F. The conceptual and statistical relationship between modularity and morphological integration. *Sys Biol* 2007; 56: 818-836.
- [69] Moore AJ, Brodie ED III, Wolf JB. Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. *Evolution* 1997; 51: 1352-1362.
- [70] Olson E, Miller R. *Morphological Integration*. Chicago, USA, University of Chicago Press 1958.
- [71] Paul A. Sexual selection and mate choice. *Int J Primatol* 2002; 23: 877-904.
- [72] Pigliucci M. Genetic variance-covariance matrices: A critique of the evolutionary quantitative genetics research program. *Biol Philos* 2006; 21: 1-23.
- [73] Phillips PC, Arnold SJ. Visualizing multivariate selection. *Evolution* 1989; 43: 1209-1222.
- [74] Phillips PC, McGuigan KL. Evolution of genetic variance-covariance structure. In CW Fox, JB Wolf, eds. *Evolutionary Genetics: Concepts and Case Studies*. Oxford, United Kingdom 2006. pp.
- [75] Plavcan MJ. Sexual dimorphism in primate evolution. *Am J Phys Anthropol* 2001 S33: 25-53.
- [76] Pomiankowski A, Iwasa Y, Nee S. The evolution of costly mate preferences I: Fisher and biased mutation. *Evolution* 2001; 45: 1422-1430.
- [77] Rausher MD. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution* 1992; 46: 616-626.
- [78] Reeve JP, Fairbairn DJ. Predicting the evolution of sexual size dimorphism. *J Evol Biol* 2001; 14: 244-254.
- [79] Reeve HK, Sherman PW. Adaptation and the goals of evolutionary research. *Q Rev Biol* 1993; 68: 1-32.
- [80] Rice SH. *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sunderland, USA, Sinauer Press 2004a.
- [81] Rice SH. Developmental associations between traits: Covariance and beyond. *Genetics* 2004; 166: 513-526.
- [82] Roff DA. *Evolutionary Quantitative Genetics*. London, United Kingdom, Chapman and Hall 1997.
- [83] Roff DA. *Life History Evolution*. Sunderland, USA, Sinauer Press, 2002.
- [84] Roff DA. The evolution of genetic architecture. In: M Pigliucci, K Preston, eds. *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes*. Oxford, United Kingdom 2004. pp. 345-365.

- [85] Rolian C. Integration and evolvability in primate hands and feet. *Evol Biol* 2009; 36: 109-117.
- [86] Setchell JM. Do female mandrills (*Mandrillus sphinx*) prefer brightly coloured males? *Int J Primatol* 2005; 26: 715-735.
- [87] Sih A, Bell A, Johnson JC. Behavioral syndromes: An ecological and evolutionary overview. *Trends Ecol Evol* 19: 372-378.
- [88] Stearns SC. *The Evolution of Life Histories*. Oxford, United Kingdom, Oxford University Press, 1992.
- [89] Strier KB. *Primate Behavioral Ecology*. Boston, USA, Allyn and Bacon, 2006.
- [90] Strier KB, Alberts S, Wright PC, Altmann J, Zeitlyn D. Primate life history databank: Setting the agenda. *Evol Anthropol* 2006; 2: 44-46.
- [91] Strier KB, Altmann J, Brockman DK, Bronikowski AM, Cords M, Fedigan LM, Lapp H, Liu X, Morris WF, Pusey AE, Stoinski TS, Alberts SC. The primate life history database: A unique shared ecological data resource. *Methods Ecol Evol* 2010; 1: 199-211.
- [92] van Schaik CP, Kappeler PM. The social systems of gregarious lemurs: Lack of convergence with anthropoids due to evolutionary disequilibrium? *Ethology* 1996; 102: 915-941.
- [93] Wade MJ, Kalisz S. The causes of natural selection. *Evolution* 1990; 44: 1947-1955.
- [94] Walker JA. A general model of functional constraints on phenotypic evolution. *Am Nat* 2007; 170: 681-689.
- [95] Walsh B. Lecture notes, short course in evolutionary quantitative genetics. 2006; <http://nitro.biosci.arizona.edu/workshops/Aarhus2006/notes.html>.
- [96] Wilson AJ, Reale D, Clements MN, Morrissey MM, Postma E, Walling CA, Kruuk LEB, Nussey DH. 2009. An ecologist's guide to the animal model. *J Anim Ecol* 79: 13-26.
- [97] Wolf JB, Brodie ED III, Cheverud JM, Moore AJ, Wade MJ. Evolutionary consequences of indirect genetic effects. *Trends Ecol Evol* 1998; 13: 64-69.
- [98] Wolf JB. Genetic architecture and evolutionary constraint when the environment contains genes. *Proc Nat Acad USA* 2003; 100: 4655-4660.