Translating Between Microevolutionary Process and Macroevolutionary Patterns: The Correlation Structure of Interspecific Data

Author(s): Thomas F. Hansen and Emilia P. Martins

Source: *Evolution,* Vol. 50, No. 4 (Aug., 1996), pp. 1404-1417

Published by: Society for the Study of Evolution

Stable URL: https://www.jstor.org/stable/2410878

Accessed: 19-01-2020 14:50 UTC

TRANSLATING BETWEEN MICROEVOLUTIONARY PROCESS AND MACROEVOLUTIONARY PATTERNS: THE CORRELATION STRUCTURE OF INTERSPECIFIC DATA

THOMAS F. HANSEN[1] AND EMÍLIA P. MARTINS[2]

[1]*University of Oslo, Division of Zoology, Department of Biology, P. O. Box 1050, Blindern, N 0316 Oslo 3 Norway*
E-mail: thomas.hansen@bio.uio.no
[2]*Department of Biology, University of Oregon, Eugene, Oregon 97403*
E-mail: emartins@work.uoregon.edu

Abstract.—As species evolve along a phylogenetic tree, we expect closely related species to retain some phenotypic similarities due to their shared evolutionary histories. The amount of expected similarity depends both on the hierarchical phylogenetic structure, and on the specific magnitude and types of evolutionary changes that accumulate during each generation. In this study, we show how models of microevolutionary change can be translated into the resulting macroevolutionary patterns. We illustrate how the structure of phenotypic covariances expected in interspecific measurements can be derived, and how this structure depends on the microevolutionary forces guiding phenotypic change at each generation. We then explore the covariance structure expected from several simple microevolutionary models of phenotypic evolution, including various combinations of random genetic drift, directional selection, stabilizing selection, and environmental change, as well as models of punctuated or burst-like evolution. We find that stabilizing selection leads to patterns of exponential decrease of between species covariance with phylogenetic distance. This is different from the usual linear patterns of decrease assumed in most comparative and systematic methods. Nevertheless, linear patterns of decrease can result from many processes in addition to random genetic drift, such as directional and fluctuating selection as well as modes of punctuated change. Our framework can be used to develop methods for (1) phylogenetic reconstruction; (2) inference of the evolutionary process from comparative data; and (3) conducting or evaluating statistical analyses of comparative data while taking phylogenetic history into account.

Key words.—Comparative method, evolution, phylogenetic analysis, population genetics, quantitative genetics, systematics.

Measurements taken from phylogenetically related species are related to each other by a complex, hierarchical structure of covariances resulting from shared evolution along a common phylogeny. The general form of this structure depends on the underlying microevolutionary processes (e.g., mutation, random genetic drift, and selection), and provides a link between those processes and the macroevolutionary pattern observed in a set of comparative or interspecific data. When the covariance structure is known, we can use comparative data to estimate the pattern of phylogenetic relationships among species, to infer the evolutionary process underlying particular characters along known phylogenies, and to conduct statistical analyses of comparative data while taking phylogenetic history into account. In the past, techniques designed to accomplish these goals either have been based on restrictive neutral models of phenotypic evolution or have not been placed within an explicit microevolutionary framework. In the current paper, we show how the covariance structure of interspecific data can be described using a minimal number of assumptions regarding the evolutionary process. This approach can be used to develop new statistical techniques or to provide a means of comparing different microevolutionary processes to one another in terms of their impact on macroevolutionary patterns of covariation.

Comparative data and a model of their covariance structure are commonly used in at least two areas of biology. In phylogeny reconstruction (see reviews by Wiley 1981; Felsenstein 1982, 1983, 1988a; Swofford and Olson 1990; Maddison and Maddison 1992), traits are chosen because their covariance structures can be related to the pattern of speci-

ation occurring in a particular clade of organisms. In maximum likelihood, distance and parsimony approaches, a model of phenotypic evolution (e.g., Brownian motion) relates a phylogenetic hypothesis to a prediction about the covariance structure in the data. The observed covariance structure can then be used to estimate the phylogeny. Distance methods, for example, can be seen as estimating a phylogeny by minimizing the difference between expected and observed covariance structures (Martins 1995).

In studies of phenotypic evolution along a phylogeny (see Brooks and McLennan 1991; Harvey and Pagel 1991; Maddison and Maddison 1992; Miles and Dunham 1993; Maddison 1994; Martins and Hansen 1996, for reviews), the process is reversed. The traits of most interest are those that are thought to have been subjected to the action of natural or sexual selection. Given a phylogeny for the clade and an expected pattern of covariation along that phylogeny, algorithms are used to estimate the location, magnitude and relationship between evolutionary changes occurring along that phylogeny or to infer the microevolutionary process underlying the characters of interest. As biologists become increasingly quantitative, it has also become clear that statistical analyses of interspecific data require implicit or explicit reference to the underlying covariance structure produced by phylogenetic relationships among the measured species. Most parametric statistics assume that there is no covariance among data points and that the variances of such data are equal (i.e., homoscedastic). Alternative statistical techniques (e.g., spatial autocorrelation, generalized least squares) have been developed for use when these assumptions are not met, but

require explicit assumptions about the underlying expected covariance structure of the comparative data.

The macroevolutionary covariance structure in all of the above cases is produced by the combination of microevolutionary forces acting on the traits being considered and the branching structure of the underlying phylogeny. In many situations, simple algorithms or statistical procedures (e.g., parsimony, spatial autocorrelation statistics) have been applied in lieu of microevolutionary models of phenotypic evolution to describe the predicted covariance structure. These procedures can be useful when little is known about the underlying evolutionary processes. Methods based on known mechanisms of genetic change (Felsenstein 1985, 1988a,b) are usually much easier to interpret in evolutionary terms. To date, Brownian motion is the only mathematical model that has been used widely to describe the covariance structure expected among continuous traits measured in different species based on explicit population genetic arguments (e.g., Edwards and Cavalli-Sforza 1964; Felsenstein 1973, 1981, 1985; Martins 1994). Although Brownian motion is a reasonable model of undirected change, it may not adequately describe other likely scenarios in which selection has also acted on the traits of interest (Felsenstein 1988b).

In this paper, we present a framework for deriving mathematical formulas to describe the macroevolutionary patterns of covariance expected among traits measured in groups of extant species. This framework requires only a few minimal assumptions regarding the type of phenotypic microevolutionary changes that are possible, and is sufficiently general to describe the evolution of both continuous and categorical characters undergoing most types of evolutionary change. We then illustrate the framework by showing how it can be used to determine the covariance pattern expected under several familiar types of gradual phenotypic evolution: (1) mutation and random genetic drift alone; (2) constant or fluctuating directional selection; and (3) stabilizing selection with random genetic drift or environmental fluctuations. We also consider models of "punctuated" or "burst-like" evolution. Although the particular results of our analyses may be of some interest in particular applications, they serve mainly as illustrations of the general approach and were chosen for simplicity as well as realism. These examples show that it is feasible to compute phylogenetic correlations from a wide range of microevolutionary models. Our approach may thus be used for simple comparisons of different comparative methods in the common currency of the expected macroevolutionary covariance structure, and aid in the development of statistical techniques derived from alternative microevolutionary models.

A MODEL OF THE INTERSPECIFIC CORRELATION STRUCTURE

Phenotypic Evolution along a Phylogenetic Tree

During each generation, forces such as selection, mutation, random genetic drift, inheritance and environmental fluctuations combine to produce evolutionary changes in the phenotype of a population. In very simple terms, phenotypic evolution can be described as a string of phenotypes that exists over stretches of evolutionary time. Because the microevolutionary forces acting in each generation have both deterministic and random elements, we can model the phenotype mathematically as a random variable, X, and the series of phenotypic states existing through time as a stochastic process, $X(t)$. If we assume that the phenotype in any single generation, $X(t)$, is a function only of the phenotype during the previous generation, $X(t - 1)$, the process has the Markov property that if $X(t - 1)$ is known, the states in previous generations, $X(t - 2)$, $X(t - 3)$, and so on, do not contain any additional predictive information about $X(t)$.

When considering comparative data, this evolutionary process unfolds along the branches of a phylogenetic tree. The process begins with a single ancestor, and results in the phenotypes of each of the extant species at the tips of the tree. To model this, we envision evolution along each branch of a phylogeny as a separate stochastic process (Fig. 1). Thus, the process starts at the root of the tree with some initial value $X_0(t_0)$, evolves along the stem according to the process $X_0(t)$ and reaches some value $X_0(t_1)$ at the first branch point. There, two new processes $X_1(t)$ and $X_2(t)$ start along the descendant branches 1 and 2, both with the initial condition $X_0(t_1)$. The processes develop along the tree in this way until the endpoints, the extant species, are reached. In this appproach, comparative or interspecific data may be viewed as a snapshot of the process at a single point in time (Fig. 1). Because the time spans between nodal points on a phylogeny usually encompass thousands of generations, we consider only continuous time models. Hence, our framework describes comparative data as the result of a continuous time, Markovian stochastic process, $X(t)$, modeling the phenotypes of species as they change through evolutionary time.

To apply this framework, the types of evolutionary changes occurring in $X(t)$ must be specified. In our simple applications below, for example, we assume that the evolving phenotype can be described in terms of its population mean value, thereby ignoring changes in other aspects of the population distribution such as the variance. We also assume that species evolve independently after they split at the nodal points on the tree, thereby limiting our study of phenotypic correlation to correlations produced by common ancestry. None of these assumptions is critical to the basic approach and can be eliminated by explicitly specifying alternative possibilities within our general approach.

The Covariance Structure Given a Model and a Tree

As phenotypic evolution unfolds along a phylogeny, evolutionary changes accumulate. The resulting taxon phenotypes are functions of (1) the initial or starting phenotype at the root of the tree; (2) the time that the taxon has been evolving; and (3) the magnitude and types of evolutionary changes that have occurred. We can think of any observed taxon phenotype as a single manifestation of a general evolutionary process that could have resulted in any one of a number of other possible taxon phenotypes. Thus, each observed phenotype can be viewed as a sample from a theoretical population of possible phenotypes, such that each taxon phenotype has several statistical properties (e.g., a mean and variance) associated with it.

The expected statistical properties of any single phenotype depend only on the direct history of the taxon from the pres-
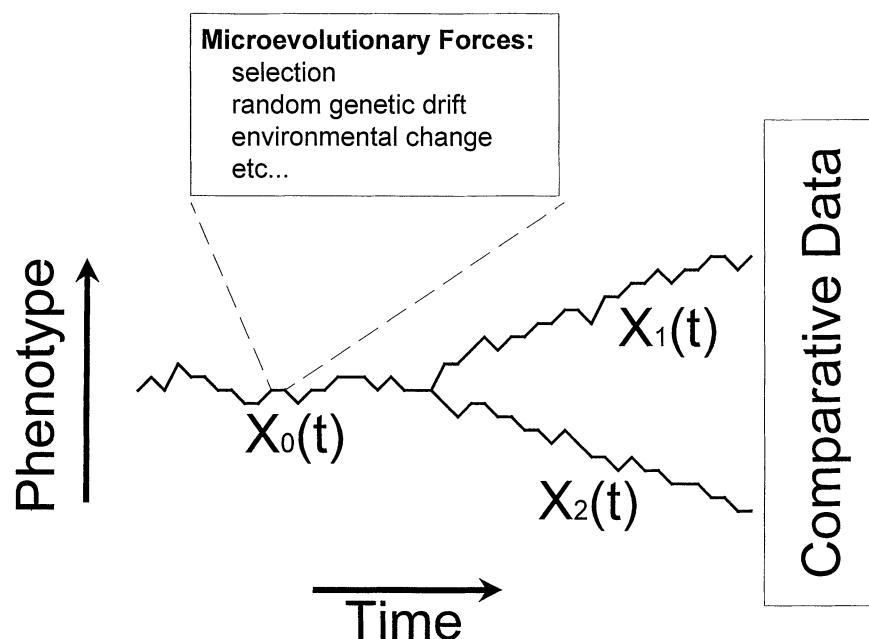
FIG. 1. An illustration of our basic model for translating between microevolutionary forces acting on a short time scale and the macroevolutionary pattern observed in a set of comparative data.

ent time to the root of the phylogeny, and can be computed without regard to the rest of the phylogeny. However, the phenotypes of two species that are phylogenetically related to one another also have an expected covariance due to the time that the two species evolved together as a single ancestor. Thus, comparative data consisting of measurements of several different species are expected to be related to one another by the complex correlation structure inherent in the phylogeny. The complete set of expected variances and covariances for all pairs of species measured in a comparative study is the overall covariance structure that is used to reconstruct phylogenetic relationships or to infer the evolutionary histories of traits along a known phylogeny.

As shown in Appendix A, if phenotypic evolution follows a Markovian process along a phylogeny with characters evolving independently along each branch, then the theoretical covariance between a single trait measured in two different species is:

$$\text{Cov}[X_i, X_j] = \text{Cov}[E[X_i|X_z], E[X_j|X_z]], \qquad (1a)$$

where X_z is the state of the character in the most recent common ancestor, z, of species i and j, and $E[X_i|X_z]$ is the expected value of X_i given X_z (i.e., the regression of X_i on X_z). Phenotypes are usually made up of more than one character, and we can use a multivariate form of eqn. (1) to describe the correlation structure of a set of mean phenotypes by considering $\mathbf{X}_i$, a column vector of character mean phenotypes in species i, rather than the single character measurement, X_i:

$$\text{Cov}[\mathbf{X}_i, \mathbf{X}_j] = \text{Cov}[E]\mathbf{X}_i|\mathbf{X}_z], E[\mathbf{X}_j|\mathbf{X}_z]]. \qquad (1b)$$

(Bold face denotes a vector or a matrix throughout this paper.) In this case, the "covariance" between the two species' trait vectors is a matrix of covariances between all possible com-

binations of traits, one from each species. Specifically, the entry in the kth row and lth column of this matrix is $\text{Cov}[X_{ik}, X_{jl}] = \text{Cov}[[E[X_{ik}|\mathbf{X}_z], E[X_{jl}|\mathbf{X}_z]]$ where X_{ik} is the measurement of the kth trait in species i and X_{jl} is the measurement of the lth trait in species j.

Equation (1) shows that the expected covariance between measurements of a set of traits from two different species may be replaced generally with the covariance of the regressions of those traits on the state of the most recent common ancestor. This reduces the problem of calculating the covariances between traits and species on a phylogeny to a problem of calculating regressions of descendants on ancestors in an evolutionary sequence. These regressions will depend on the details of the microevolutionary forces acting within lineages, and eqn. (1) thus provides a link between microevolutionary processes and the resulting macroevolutionary structure.

Equation (1) specifies only the covariance due to common ancestry. If the species do not evolve independently, a term specifying the correlated evolution must be added to the covariance (Appendix A). Some further remarks on this point are given in the discussion.

COVARIANCE STRUCTURE EXPECTED FROM SOME SIMPLE MICROEVOLUTIONARY MODELS

In this section, we illustrate the use of our approach by deriving the pattern of phylogenetic covariance in quantitative traits expected from several simple models of microevolutionary process. The main results from these models are summarized in Table 1 and Figure 2.

Models of Neutral Phenotypic Evolution

Models that describe the gradual evolution of quantitative characters under mutation and random genetic drift alone are

TABLE 1. A summary of some phenotypic covariance structures expected under microevolutionary models discussed in the text. Variables are defined in the main text, and some of the formulas are simplified for clarity.

Model	$\text{Cov}[\mathbf{X}_i, \mathbf{X}_j] \approx$	Correlation between species phenotypes
Random genetic drift:		
alone	$(\mathbf{G}/N_e)t_z$	Decreases linearly with increasing phylogenetic distance
drift-mutation balance	$2\mathbf{G}_m t_z$	(but may take other forms)
Directional selection:		
with random genetic drift	$(\mathbf{G}/N_e)t_z$	Decreases linearly with increasing phylogenetic distance.
fluctuating environment	$\mathbf{G}\mathbf{V}_s\mathbf{G}'t_z$	Trend cannot be detected from comparative data alone.
Stabilizing selection and drift:		
univariate	$\hat{V}\text{Exp}[-wV_A t_{ij}]$	Decreases exponentially with increasing phylogenetic distance.
multivariate	$\mathbf{Q}(t_{iz})\hat{\mathbf{V}}\mathbf{Q}'(t_{jz})$	Decreases with increasing phylogenetic distance as a sum of exponentials.
Stabilizing selection + environmental change:		
Brownian motion of optimum	$\mathbf{E}_\theta t_z(+ \text{ exponentials})$	Decreases linearly with increasing phylogenetic distance.
Lynch-Lande model	$\mathbf{Q}(t_{iz})\mathbf{E}_\theta \mathbf{Q}'(t_{jz})$	Decreases with increasing phylogenetic distance as a sum of exponentials.
Punctuated phenotypic change	$\left(\mathbf{H} + \dfrac{\sigma^2}{\mu}\mathbf{h}\mathbf{h}'\right)\dfrac{t_z}{\mu}$	Decreases linearly with increasing phylogenetic distance.
Change correlated with speciation events	$\mathbf{H}N_z$	Decreases with ratio of shared to total speciation events.

frequently used as null models of the evolutionary process (e.g., Edwards and Cavalli-Sforza 1964; Felsenstein 1973, 1981; Lande 1977; Turelli et al. 1988; Lynch 1989, 1993; Spicer 1993; Cheetham et al. 1993, 1994; Martins 1994). To begin, consider a group of organisms exhibiting a set of traits that are assumed to be the sum of additive genetic and environmental or residual components. As in most quantitative genetic applications, we further assume that both additive genetic and environmental components have multivariate
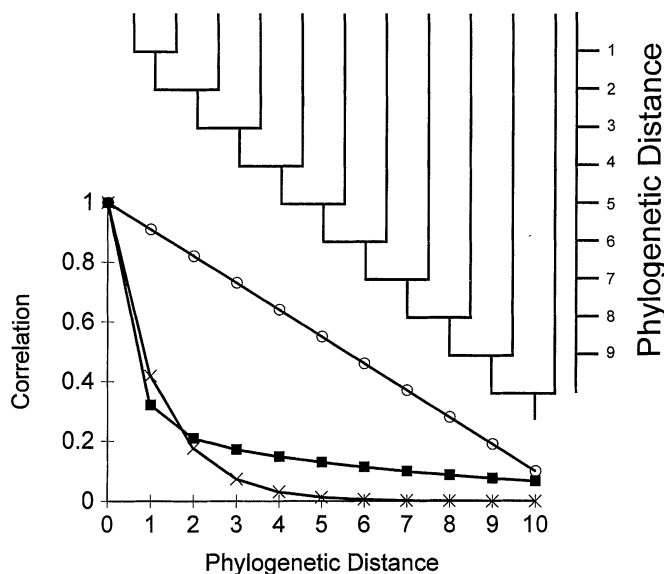

FIG. 2. The relationship between phylogenetic distance and the correlation between species phenotypes. Circles illustrate the linear decrease in correlation with phylogenetic distance expected under a Brownian motion model of random genetic drift, directional selection, or stabilizing selection with shifting optimum. Crosses refer to the exponential decrease in correlation expected for a single character undergoing stabilizing selection. Squares refer to decrease of correlation as the sum of exponentials expected for several characters undergoing stabilizing selection.

normal distributions, that the two types of components are independent of one another, and that the environmental component has a mean of zero. These assumptions imply the absence of dominance and epistatic components of variance as well as genotype-environment interactions and correlations, but are still useful for illustrative purposes.

Let $\mathbf{X}(t)$ be the mean phenotype of the population at time t, and $\mathbf{G}$ be the additive genetic variance-covariance matrix which is assumed to be constant from generation to generation (Fig. 3). If gametes are sampled at random from the current generation to produce the set of individuals in the next generation, then the mean phenotype in the next generation $\mathbf{X}(t + 1)$ is normally distributed with mean $\mathbf{X}(t)$ and variance-covariance $\mathbf{G}/N_e$, where N_e is the (variance) effective population size. Thus, the change in the mean phenotype during each generation is $\Delta \mathbf{X} = \mathbf{X}(t + 1) - \mathbf{X}(t) \sim N(\mathbf{0}, \mathbf{G}/N_e)$, where the notation $\sim N(\mathbf{a},\mathbf{b})$ means that a variable is normally distributed with mean vector $\mathbf{a}$ and variance-covariance matrix $\mathbf{b}$. Because sampling of gametes in one generation is independent of sampling in the next generation, the mean phenotype will also be normally distributed with constant mean and variance proportional to the amount of time that has passed $(\mathbf{X}(t) \sim N(\mathbf{X}(0), (\mathbf{G}/N_e)t))$. By definition, the mean phenotype evolves as if by Brownian motion, a stochastic process in which the changes occurring at each interval are independent, normally distributed random variables with variance-covariance proportional to the time interval (e.g., Karlin and Taylor 1975).

According to eqn. (1), the expected covariance between the mean phenotypes of two species depends on the regressions of their mean phenotypes on the mean phenotype of their most recent common ancestor. In the case of Brownian motion, the expected value of a species mean phenotype at any point in time, $\mathbf{X}(t)$, given the mean phenotype of an ancestor living at time t_z, $\mathbf{X}(t_z)$, is the mean phenotype of that ancestor, $E[\mathbf{X}(t)|\mathbf{X}(t_z)] = \mathbf{X}(t_z)$. Using this result with eqn. (1), we find that $\text{Cov}[\mathbf{X}_i, \mathbf{X}_j] = \text{Cov}[E[\mathbf{X}_i|\mathbf{X}_z]], E[\mathbf{X}_j|\mathbf{X}_z]] = \text{Cov}[\mathbf{X}_z, \mathbf{X}_z] = \mathbf{V}(t_z)$. Hence, the covariance of traits in
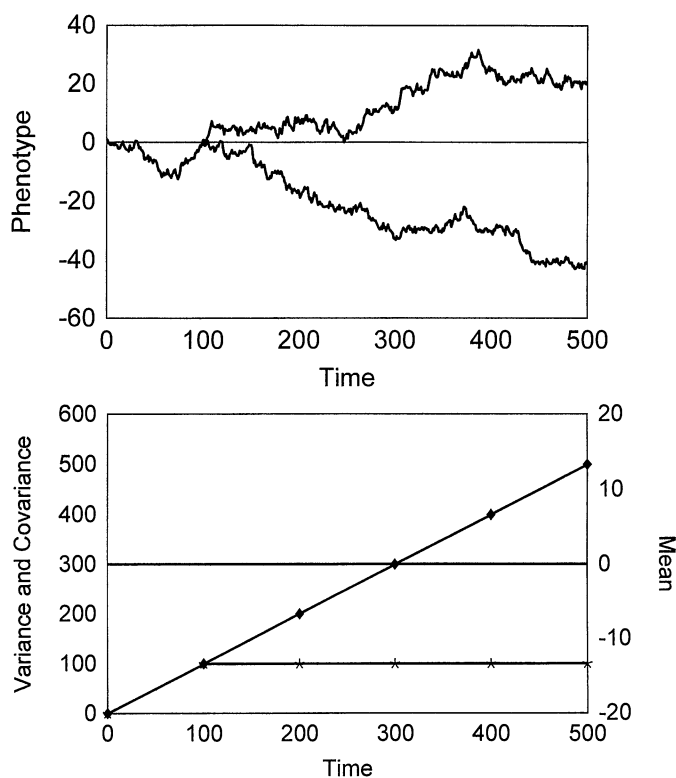
Random Genetic Drift

FIG. 3. An illustration of phenotypic change expected under a Brownian motion model of phenotypic evolution under random genetic drift. Brownian motion can also be used to describe phenotypic evolution under fluctuating, but on average undirected, directional selection. Diamonds show changes in the variance through phylogenetic time, asterisks denote changes in the covariance, and the solid line denotes shifts in the expectation of the phenotype.

two species expected under this model of neutral evolution equals the variance-covariance of the most recent common ancestor of those two species. The variance-covariance of the ancestor, $V(t_z)$, is computed in Appendix B as $V(t_z) = V_0 + (G/N_e)t_z$, where V_0 is the variance at the root of the tree. As we are interested in the relative similarity of species within a clade, we standardize by conditioning on the state at the root of the tree, thereby setting $V_0 = 0$. Then, the covariance between measures of any two species becomes:

$$\text{Cov}[X_i, X_j] = \frac{G}{N_e} t_z. \qquad (2)$$

Hence, the covariance between the two species phenotypes is proportional to the time the species evolved together as one common ancestor. If the two species are extant, then $t_z = t - t_{ij}/2$, where t is the time from the root of the tree until present and t_{ij} is the time separating the two species on the tree. Thus, the covariance decreases linearly with the time separating the two species (Fig. 2).

Felsenstein (1988b) and Lynch (1989) elaborated on this model by building on the results of Lynch and Hill (1986) and assuming that the genetic variance is in (deterministic) mutation-drift equilibrium. In this second model of neutral

phenotypic evolution, the expected genetic variance-covariance in generation t + 1 is equal to the genetic variance-covariance in the previous generation minus the expected amount of variation lost by random genetic drift, $G/2N_e$, plus the amount of new variation generated by mutation, G_m. Symbolically, $G(t + 1) = G(t)(1 - 1/(2N_e)) + G_m$. Solving this yields an equilibrium expected genetic-variance covariance, $G = 2N_e G_m$. Combining this with eqn. (2) yields:

$$\text{Cov}[X_i, X_j] = 2G_m t_z. \qquad (3)$$

For a single character, eqn. (3) becomes $\text{Cov}[X_i, X_j] = 2V_m t_z$, where V_m is the well known mutational variance parameter (for discussion and estimates, see Lynch 1988).

The results presented in equations (2) and (3) are well known. They are multivariate versions of the model used by Edwards and Cavalli-Sforza (1964) and Felsenstein (1973, 1981) to develop methods for phylogeny reconstruction; by Felsenstein (1985) to develop his method of independent contrasts for the analysis of comparative data; by Lande (1977), Turelli et al. (1988), Lynch (1989), Spicer (1993), and Cheetham et al. (1993, 1994) to test null hypotheses of random genetic drift in comparative data; and by Martins (1994) as a means of estimating V_m from comparative data.

There are several incomplete features of the Brownian motion models described above as reasonable models of phenotypic evolution under random genetic drift. First, although the additive genetic variance-covariance matrix, G, is likely to change stochastically through evolutionary time, we have considered only the case in which G is in deterministic equilibrium. This may be reasonable if the evolving phenotypes are controlled by large numbers of independent loci, such that sampling effects on the variance are small. Otherwise, a model which incorporates the effects of sampling variation in G (e.g., Lynch and Hill 1986) may be preferred. Second, the Brownian motion models assume that mutations are unbiased and independent of the state of the character. Alternative models of the mutation process may lead to mathematical models of phenotypic evolution which differ considerably from Brownian motion. For example, Cockerham and Tachida (1987) presented an alternative model of neutral evolution based on the assumption that the allelic state of a new mutation is independent of the previous allelic states of the locus. In this model the phenotypic evolutionary changes produced by new mutations are not independent of earlier allelic states of the locus. When the locus reaches extreme values, evolutionary changes shift the phenotype away from those extremes towards an average value determined by the mutational distribution. Evolution of a phenotype determined by several such loci is not well described by Brownian motion. A model such as the Ornstein-Uhlenbeck (OU) process we consider below, which includes a centripetal force that gets successively stronger as the phenotype gets further and further away from the mutational average, would be more appropriate. As discussed below, the OU model leads to a very different covariance structure than does Brownian motion.

Models Incorporating Directional Selection

The phenotypic response to constant or fluctuating directional selection can be modeled as a Brownian motion process
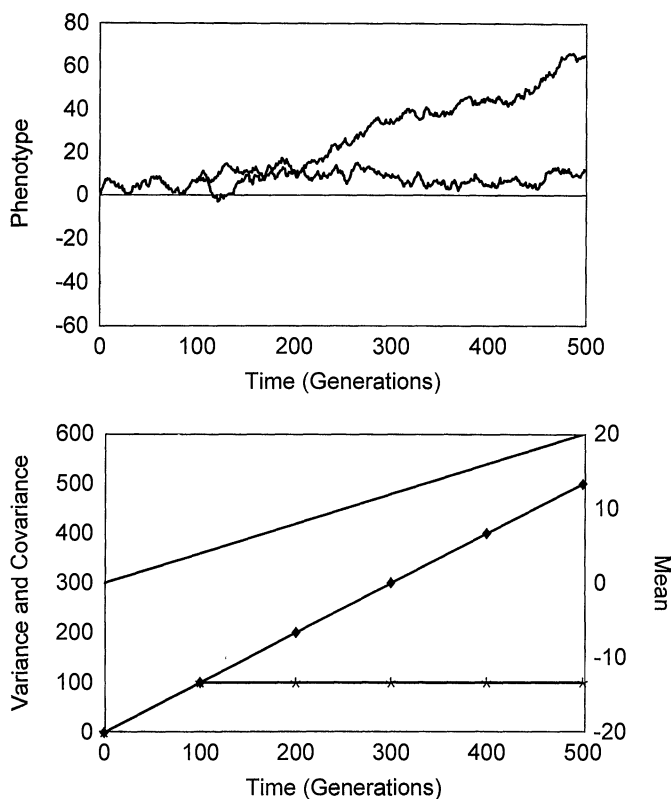
Directional Selection

FIG. 4. An illustration of phenotypic change expected under a model of directional selection (Brownian motion with a trend). The random fluctuations can be due to random genetic drift or random fluctuations in the selection differential. Diamonds show changes in the variance through phylogenetic time, asterisks denote changes in the covariance, and the solid line denotes shifts in the expectation of the phenotype.

with a trend. To the model of neutral evolution in the previous section, we add directional selection through the linear fitness function $w(\mathbf{x}) = \mathbf{s}'\mathbf{x}$, where $\mathbf{x}$ is a column vector of individual phenotypes and $\mathbf{s}'$ a row vector of selection differentials (prime denotes transposition). In this case, the change in the mean phenotype, $\mathbf{X}$, over a single generation, under first selection and then random sampling of gametes, has expectation $\mathbf{Gs}$ and variance $\mathbf{G}/N_e$,where again $\mathbf{G}$ is the additive genetic variance-covariance matrix, and N_e is the effective population size. On long time scales, the evolution of $\mathbf{X}$ may be described by a diffusion process with infinitesimal mean vector (i.e. expected change per time over short time intervals) $\mathbf{Gs}$, and infinitesimal variance matrix (i.e. variance of change per time over short time intervals) $\mathbf{G}/N_e$. This results in a Brownian motion process with a trend (Fig. 4).

Alternatively, if the strength and direction of selection fluctuate, we might (as discussed in Felsenstein 1973, 1988b) describe the strength of selection, $\mathbf{s}$, as a random variable drawn independently at short intervals of time from a normal distribution with mean $\mathbf{m}_s$ and variance, $\mathbf{V}_s$, such that $\mathbf{s}(t) \sim N(\mathbf{m}_s, \mathbf{V}_s)$.This model describes the selection differential fluctuating in an uncorrelated manner, as a white noise process, and results in a diffusion process with infinitesimal

mean $\mathbf{Gm}_s$ and infinitesimal variance $\mathbf{G}/N_e + \mathbf{GV}_s\mathbf{G}$. Again, this is a Brownian motion process with a trend. To apply this model, we need only assume that environmental fluctuations are independent of one another on a phylogenetic time scale (e.g., over millions of generations). Correlations between environmental changes occurring over thousands or tens of thousands of generations will probably not violate the assumptions of this model.

The trends found in the above models result in an overall shift of the resulting species phenotypes in one direction or another, but do not affect the relationships among species phenotypes (i.e., the covariance structure, see Fig. 4). In all cases, the expected covariance between species phenotypes is identical to that found for the simple Brownian motion models of neutral phenotypic evolution (Fig. 2). Also, as with the neutral models, the covariance between species is expected to depend linearly on time such that species that are more distantly related to one another are expected to be less similar to one another than are closely related species. Thus, comparative and systematic methods that assume this linear decrease in covariance with phylogenetic distance (e.g., Felsenstein 1985) can be applied with models of both neutral and directed phenotypic evolution.

If, however, statistical methods are being used to obtain estimates of microevolutionary processes from interspecific data, models of directional selection can lead to different interpretations of the estimated parameters than the models of neutral evolution. Whereas the covariance between species evolving by the above model of constant directional selection is identical to that found for species evolving by a model of neutral evolution, the model of fluctuating selection leads to: $Cov[\mathbf{X}_i, \mathbf{X}_j] = Var[\mathbf{X}_z] = (\mathbf{G}/N_e + \mathbf{GV}_s\mathbf{G})t_z$. If the variance due to selection is much greater than that due to random genetic drift, then:

$$Cov[\mathbf{X}_i, \mathbf{X}_j] \approx \mathbf{GV}_s\mathbf{G}t_z. \qquad (4)$$

Thus, for example, procedures developed to estimate elements of the $\mathbf{G}$ matrix from interspecific data will differ fundamentally depending on whether the model of evolution by random genetic drift or by fluctuating selection is assumed.

Models of Stabilizing Selection and Random Genetic Drift

Lande (1976a, 1979) developed a model for the evolution of quantitative characters under stabilizing selection and random genetic drift, based on the hypothesis (Lande 1976b) that a balance between mutation and selection keeps the genetic variance of the trait constant. As in the previous examples, this model begins by considering a population of organisms with phenotypes that are the sum of additive genetic and environmental components. Both the genetic and environmental components are multivariate normals, independent of each other, and the environmental component has a mean of zero. Selection operates according to a Gaussian fitness function $w(\mathbf{x}) = \exp[-(\mathbf{x} - \theta)'\mathbf{W}(\mathbf{x} - \theta)/2]$, where θ is a column vector containing the optimum values of the traits, $\mathbf{W}$ is a symmetric positive definite matrix of selection parameters, and a prime denotes transposition. The elements of $\mathbf{W}$ may be interpreted as partial regression coefficients of

fitness on quadratic deviations of the traits from their optima, so that w_{ij} is the partial regression of fitness on $(x_i - \theta_i)(x_j - \theta_j)$. The diagonal elements, w_{ii}, measure the amount of direct stabilizing selection on trait i, and the assumption of positive definiteness implies that all characters are under stabilizing selection.

Using this model, Lande (1979) showed that in a finite population, with random sampling of gametes occurring after selection, the expected change in the mean phenotype in one generation is $-G(W^{-1} + P)^{-1}(X - \theta) \approx -GW(X - \theta)$, where G is the additive genetic variance-covariance matrix and P the phenotypic variance-covariance matrix. The approximation $G(W^{-1} + P)^{-1} \approx GW$, which we shall use throughout, assumes that selection is weak (i.e., $W << P$). Lande (1979) also showed that the variance of change in one generation is G/N_e. Hence, if G, P, W, θ, and N_e remain constant over long stretches of evolutionary time, $X(t)$ defines a multivariate Ornstein-Uhlenbeck (OU) process, a diffusion process with infinitesimal mean $-GW(X - \theta)$ and infinitesimal variance G/N_e. For a single character, the model becomes a univariate OU process with infinitesimal mean $-wV_A(X - \theta)$, and infinitesimal variance V_A/N_e, where V_A is the additive genetic variance and w the strength of stabilizing selection (Lande 1976a).

Ornstein-Uhlenbeck processes are the "rubber-band" processes that have often been proposed as ways of modeling the phenotypic response to stabilizing selection (e.g., Felsenstein 1988b; Martins 1994). In these processes, the mean phenotypes evolve in a driftlike stochastic process, but are pulled towards the central optimum, θ, by a restraining force (Fig. 5). In the above case, the strength of that force is proportional to GW, and increases as the phenotype wanders further from the optimum, as if tied to the optimum by a rubber-band. Brownian motion can be considered to be a restricted form of an OU process in which the strength of the restraining force is zero.

In Appendix B, we show that for an OU process, the regression of a species phenotype on its ancestor is $E[X(t)|X(t_z)] = Q(t_{zi})X(t_z) + (I - Q(t_{zi}))\theta$, where $Q(t) = \exp[- GWt]$, I is the identity matrix, and t_{zi} is the time separating species i from its ancestor, z. Also as shown in Appendix B, the variance in $X(t)$ is $V(t) = \hat{V} - Q(t)\hat{V}Q'(t)$, where $\hat{V} = (2N_eW)^{-1}$ is the equilibrium, or stationary, variance of the process, and the variance at the root of the tree has been set to zero. In Appendix C, we show that under Gaussian processes, the entire distribution of character states for any points on the tree is multivariate normal. Thus, the character state vector at any single point is completely defined by the mean vector and variance-covariance matrix given above. Using eqn. (1), the covariance structure can be computed as:

$$Cov[X_i, X_j] = Q(t_{zi})\hat{V}Q'(t_{zj}) - Q(t)\hat{V}Q'(t) \quad (5)$$

where the second term goes to zero when the process reaches stationarity (i.e., when all influence of the phenotypes at the root of the tree has been lost). For a single character, the covariance between two species phenotypes is:

$$Cov[X_i, X_j] = \hat{V}(\exp[-V_Awt_{ij}] - \exp[-2V_Awt]). \quad (6)$$

The correlation between species phenotypes is obtained
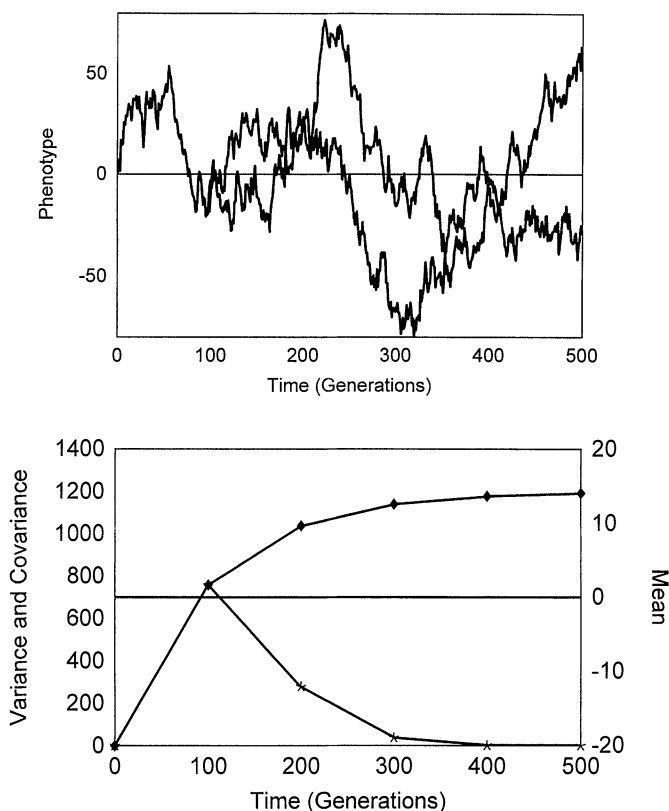
Stabilizing Selection

FIG. 5. An illustration of phenotypic change expected under stabilizing selection with random fluctuations due to random genetic drift or short scale fluctuations of the optimum (an Ornstein-Uhlenbeck process). In this illustration, the phenotype is pulled towards an (average) optimum of zero. Diamonds show changes in the variance through phylogenetic time, asterisks denote changes in the covariance, and the solid line denotes shifts in the expectation of the phenotype.

through standardizing the covariance between species by the standard deviations of the traits in each species independently. Thus, in the stationary OU case, the correlation between two species phenotypes is:

$$Corr[X_i, X_j] = \exp[-V_Awt_{ij}]. \quad (7)$$

Thus, for a single trait, the correlation between two species phenotypes depends on the strength of selection, w, the additive genetic variance, V_A, and the time separating the two species, t_{ij}. Under stabilizing selection, covariances and correlations decrease exponentially with increasing time since species divergence (Fig. 2). If the process has not reached equilibrium, the changing variance in the ancestors complicate, but do not qualitatively change, this overall picture. Similarly, correlations between species decrease more quickly if the strength of selection is increased. Closely related species are exponentially more similar to one another than are distantly related species, and this pattern is more pronounced in clades undergoing strong selection.

The multivariate case is considerably more complex as evolution of a trait is influenced by other traits due to cor-

related selection and genetic correlations. Again assuming that the process has reached equilibrium, the covariance between a trait X_i in species i and a trait X_j in species j (i.e., one element of eqn. [5]) has the general form

$$Cov[X_i, X_j] = \Sigma_r \Sigma_m c_{rm} exp[-(\lambda_r + \lambda_m)t_{ij}] \quad (8)$$

where λ denote eigenvalues of $\mathbf{GW}$, and the indices r and m extend over all eigenvalues. The coefficients c are functions of the eigenvectors of $\mathbf{GW}$ and the elements of $\hat{\mathbf{V}}$. Equation (8) shows that the covariance between species can be divided into several components with characteristic rates of decrease with phylogenetic distance. These rates of decrease are given by all pairwise sums of eigenvalues of $\mathbf{GW}$. The components are associated with different axes in phenotype space. When t_{ij} is very small, the covariance between traits equals $\hat{\mathbf{V}}$. As t_{ij} increases, components of covariance along directions in phenotype space where there is ample genetic variation and where selection is strong are rapidly lost. Among more distantly related species, only the slowly decreasing components of covariance associated with small eigenvalues are left. These are associated with directions in phenotype space along which there is little genetic variation and selection is weak (Fig. 2).

The main difference between OU and Brownian motion models is that distantly related species are expected to be relatively less similar to one another under models including stabilizing selection. In fact, under models of stabilizing selection, at least some of the eigenvalues of $\mathbf{GW}$ must be on the order of $1/t_{ij}$ for there to be any detectable covariance between species (i.e., a "phylogenetic effect"). As the natural time scale of the parameters is one generation and the species divergence times typically are on the order of millions of generations, $1/t_{ij}$ may easily be on the order of 10^{-6} or more. Only if the characters are constrained genetically in the sense that some eigenvalues of $\mathbf{G}$ are on the order of $1/t_{ij}$ or if selection is very weak in the sense that some eigenvalues of $\mathbf{W}$ are on the order of $1/t_{ij}$ will there be any detectable phylogenetic covariance possible under this model. Also note that selection must not be much stronger than random genetic drift (i.e., $\mathbf{W}$ must be on the order of $1/N_e$ or less) for there to be appreciable variance among species in the clade of interest. Stronger selection or smaller effective population sizes leads to the loss of all phenotypic variation among species.

Models of Stabilizing Selection and Environmental Change

When traits exhibit extensive variation across species, it may seem unrealistic to assume that the optimum phenotypes have been constant throughout the history of the clade. More realistically, we might model the evolution of phenotypes subjected to stabilizing selection in which the optima also change through time. In such cases, the relationships among species phenotypes are determined by the environmental processes changing the optima as well as by the direct process of evolutionary adaptation. We can model this situation by allowing the optimum phenotype to vary according to some specified stochastic process. For example, suppose that the optimum phenotypes in the OU model described above change according to a Brownian motion process. The co-

variances between species mean phenotypes under this compound model is a mixture of correlations generated by (1) the changing optima; (2) stabilizing selection and random genetic drift as discussed above; and (3) the interaction of these two forces (Hansen and Martins, unpubl.). If the force of stabilizing selection is not very weak (on the order of $1/t_{ij}$ as discussed above), the covariance among species phenotypes will be dominated by shifts in the optima rather than by random genetic drift along the phylogeny. In analogy to the results found for the neutral models above, the covariance between the phenotypes of species i and j is

$$Cov[\mathbf{X}_i, \mathbf{X}_j] \approx \mathbf{E}_\theta t_z \quad (9)$$

where $\mathbf{E}_\theta$ is the variance-covariance of change in the optima, and t_z is the time from the root of the tree to the most recent common ancestor, z, of the two species.

Lynch and Lande (1993) modeled environmental change as a white noise process (with a deterministic trend which we ignore here) affecting evolution of phenotypes evolving under stabilizing selection and random genetic drift. In a multivariate extension of their model, we let the environmental optimum, θ, in the above model of stabilizing selection be a set of serially uncorrelated, normally distributed random variables with mean vector $\mathbf{0}$ and variance-covariance $\mathbf{E}_\theta$. In this model, the infinitesimal mean and variance of change in mean phenotypes are $-\mathbf{GWX}$ and $\mathbf{GWE}_\theta + \mathbf{G}/N_e$. If the stochasticity due to random genetic drift, $\mathbf{G}/N_e$, is small enough to be negligible and the process has reached stationarity, the interspecific variance is $\hat{\mathbf{V}} = \mathbf{E}_\theta$. The covariance between species phenotypes is

$$Cov[\mathbf{X}_i, \mathbf{X}_j] = \mathbf{Q}(t_{iz})\mathbf{E}_\theta \mathbf{Q}'(t_{jz}) \quad (10)$$

where as before $\mathbf{Q}(t) = exp[-\mathbf{GW}t]$. Thus, the covariance decreases exponentially with phylogenetic distance at rates comparable to that found in the model of stabilizing selection in a constant environment.

These two models illustrate how phylogenetic correlations are strongly dependent on the mode of environmental change. Only when the environmental changes are uncorrelated on time scales shorter than those where stabilizing selection leave traces of history, as in the Lynch and Lande model (1993), is the characteristic correlation pattern of stabilizing selection apparent. If the process of environmental change leaves strong historical traces such as in the Brownian motion model above, it dominates the between species correlations and cannot easily be erased by microevolutionary forces. Phylogenetic information results from shifts in the optimum phenotypes, and is not affected by the magnitude of selective or genetic parameters. Instead, the pattern of phylogenetic correlation depends on the specific type of environmental change that is causing shifts in the optima. Unfortunately, there are few theoretical or empirical generalizations as to the modes of environmental change that might be expected. As it stands, environmental fluctuations that change the optimum phenotype might result in any one of a number of widely different interspecific covariance structures.

Models of Punctuated Evolution

The models considered thus far are based on standard quantitative genetic models of phenotypic evolution in which

small changes during each generation gradually accumulate to form major shifts in the species' mean phenotypes. Other authors, however, have argued that evolutionary change occurs in bursts at periodic intervals, alternating with long periods of phenotypic stasis (e.g., Gould and Eldredge 1993). Lande (1985, 1986) described how a model of punctuated evolution can be approximately derived by extending the above model of evolution under weak stabilizing selection with random genetic drift to a multipeaked adaptive landscape. In this model, a species mean phenotype is expected to spend most of its time on peaks of the adaptive landscape, such that it is mostly in a period of near stasis. Occasionally, the phenotype will shift from one peak to another. Under Lande's model, however, the total time spent passing between peaks is expected to be insignificant compared to the amount of time that the phenotype spends in near stasis on top of the peaks. Thus phenotypic evolution seems punctuated or "burst-like."

Bursts of change may occur as a result of the incorporation of rare favorable alleles, as a response to periods of drastic environmental change, or as major phenotypic shifts from one adaptive peak to another. If the series of peak shifts, environmental changes, or fixations of new mutations leading to bursts of phenotypic evolution occur at random points in time, the times between events (the interevent times) can be modeled as independent random variables with mean μ and variance σ^2. A random variable, $N(t)$, that counts the number of events that have occurred until time t can then be described as a renewal counting process (Karlin and Taylor 1975; Appendix D). Gillespie (1991) employed such renewal counting processes as models of protein evolution by amino acid substitution. He argued that substitutions are largely adaptive and occur in bursts when the protein is "environmentally challenged." Bursts of change were modeled by letting the variance of the interevent times, σ^2, be larger than the mean, μ, and Gillespie (1991) found that the model of clumped change fit the observed pattern better than did the standard neutral model of a molecular clock (a Poisson process in which $\sigma^2 = \mu$). The renewal counting process model we develop to describe phenotypic evolution might also be viewed as a direct extension of Gillespie's model of protein evolution by adaptive substitutions.

Under such a process, at each evolutionary event, the species' mean phenotypes, $X(t)$, may change. We assume that these evolutionary changes in phenotype, ΔX, are independent and identically distributed with mean, h, and variance-covariance matrix, H (Fig. 6). If h is zero there is no net directional evolutionary change. If h is different from zero, a long-term directional trend appears. The state of the phenotype at time t can be written as the sum of the phenotype in an ancestor z plus the changes that have happened thereafter:

$$X_i = \sum_{k=1}^{N(t_{zi})} \Delta X_k + X_z \qquad (11)$$

where X_i is the mean phenotype of species i, ΔX_k is the change occurring during evolutionary event k, and $N(t_{zi})$ is the number of evolutionary events occurring during the time between ancestor z and species i. Because the evolutionary
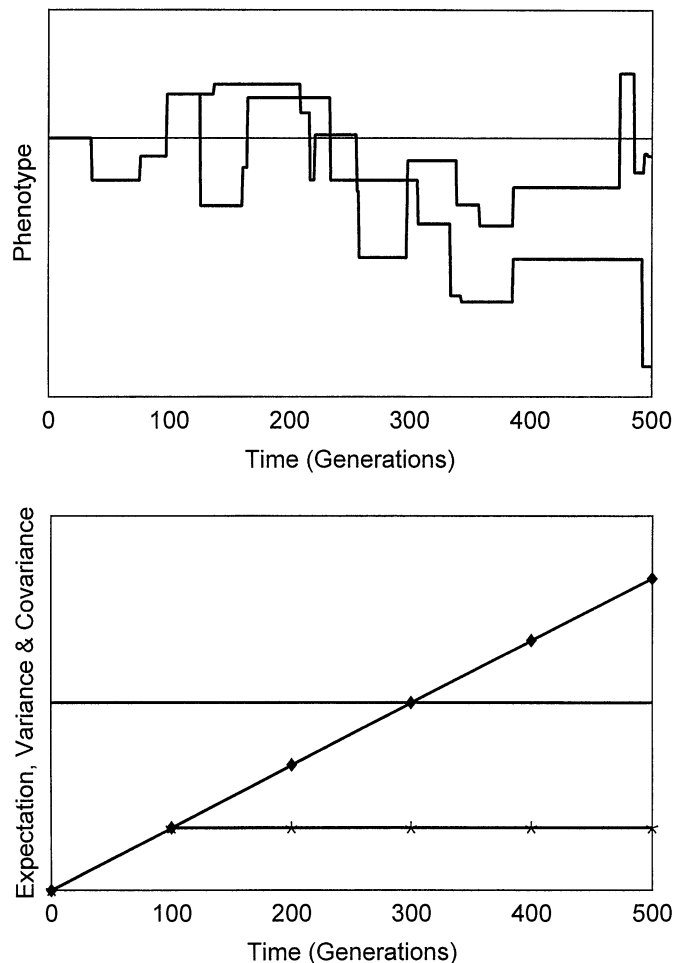
Burst-like Change

FIG. 6. An illustration of phenotypic change under a model of punctuated change occurring at random points in time. Although evolutionary change according to this model is quite different from evolutionary change under the Brownian motion model illustrated in Figure 3, changes in the means, variances and covariances of the two processes are identical. Diamonds show changes in the variance through phylogenetic time, asterisks denote changes in the covariance, and the solid line denotes shifts in the expectation of the phenotype.

changes (ΔX_k) are independent of the state of the ancestor, the expected state of the character is given by the state of the ancestor ($E[X_i|X_z] = X_z$), as it was for the Brownian motion models. Substituting for $E[X_i|X_z]$ in eqn. (1), the covariance between the mean phenotypes of species i and j is derived in Appendix D as

$$Cov[X_i, X_j] = V(t_z) = \left(H + \frac{\sigma^2}{\mu^2}hh'\right)\frac{t_z}{\mu}. \qquad (12)$$

As with the neutral and directional selection Brownian motion models, this burst-like model of phenotypic evolution yields a linear covariance structure in which the similarity between species' mean phenotypes is expected to decrease linearly with distance (Fig. 2). Thus, microevolutionary mod-

els of both gradual and punctuated evolution can lead to very similar macroevolutionary phenomena, and it may not be easy to distinguish between these two forms based only on comparative data. Note, however, that the model of punctuated change does not lead to a normal distribution of species mean phenotypes except under very special circumstances (Appendix D). Our model of punctuated evolution is also different in that trends imposed by directional selection will influence the covariance among species phenotypes. Unfortunately, without other information, it is impossible to disentangle the effects of directional trends in this model from the effects of the variance, **H**. Another important point to note is that the linear relation of covariance to phylogenetic distance resulting from our basic model of punctuated evolution (eqn. [12]) holds true regardless of the specific type of renewal process. There will only rarely be need to make any specific assumptions about the type of process (e.g., that evolution occurs via a Poisson process) in order to use this model in the analysis and interpretation of comparative data. Note, however, that eqn. (12) depends critically on the assumption that the changes at each evolutionary event are independent of each other and the times separating them. This assumption has some support from Lande's (1985, 1986) finding that the probability of a peak shift is not strongly influenced by the distance to or height of the peak to which the jump occurs. This depends on the adaptive landscape being largely uncorrelated on the level of peak positions. If the landscape is structured so that jumps toward one or a few central states are particularly likely, an exponential correlation structure more like that of an OU process may be expected.

The theory of punctuated equilibrium is associated with the idea that most or all evolution happens during speciation events (Elderedge and Gould 1972; Gould and Elderedge 1993). This is not captured in the above model, which has its underpinning in the theory of peak shifts or fixation of rare advantageous alleles, neither of which need be associated with speciation events. A model based on evolution restricted to speciation events should: (1) incorporate the known speciation events implied by the phylogeny; and (2) model the frequency of speciation events unknown due to later extinction of lineages. When all speciation events are known and phenotypic evolutionary changes associated with speciation events are independent of both the ancestral value of the character and other changes in contemporaneous lineages, then the regression of a species on an ancestor is $E[\mathbf{X}_i|\mathbf{X}_z] = \mathbf{X}_z$. If the phenotypic change at a speciation event is a random variable with mean zero and variance **H**, the covariance between two species is $Cov[\mathbf{X}_i, \mathbf{X}_j] = Var[\mathbf{X}_z] = \mathbf{H}N_z$, where N_z is the number of speciation events occurring from the root of the phylogeny to the common ancestor of the two species.

DISCUSSION

This study proposes a general approach to describe the patterns of covariation expected in comparative or interspecific data due to the evolution of species along a phylogeny. The approach is an analytical description of the translation of microevolutionary process into macroevolutionary pattern, and can be used to describe the evolution of continuous or categorical characters under a wide variety of evolutionary scenarios. We begin by describing the evolution of species phenotypes as a Markovian stochastic process. As this process unfolds along a phylogenetic tree, the phenotypes of any two phylogenetically related species are expected to be similar to some degree because of their shared evolutionary history. Measurements of an entire clade will consist of a set of phenotypes with a hierarchical covariance structure due to the shared evolutionary histories of species and the accumulation of microevolutionary changes during that shared time. We then describe the direct mathematical relationship between the form and magnitude of microevolutionary changes occurring at each generation and the resulting covariance structure of interspecific data. This covariance structure can be used to translate between microevolutionary process and macroevolutionary pattern, and as a critical first step in developing and evaluating models used both in phylogeny reconstruction and phylogenetic comparative methods.

We illustrate the use of our framework by exploring some commonly used microevolutionary models and the translation of these models into the covariance among species expected in interspecific data. By comparing the covariance structure expected from different microevolutionary processes, we can gain insight into observed macroevolutionary phenomena and explain these phenomena in explicit mathematical terms. Although we consider models of evolution under random genetic drift and stabilizing, directional, and fluctuating selection, we would like to emphasize that the models reviewed herein are presented as simple illustrations and may not be particularly realistic as a means of describing the evolution of many characters. More complex models need to be developed that incorporate several of these models acting on different parts of a phylogeny and in different parts of a multivariate phenotype.

Even the two main assumptions employed in our applications may not seem reasonable to some evolutionary biologists. These assumptions are: (1) that evolution is Markovian; and (2) that sister species evolve independently after their divergence from a common ancestor. Losos and Adler (In press) question whether speciation is well modeled as a Markovian process given the possibility of vicariance events, peripatric speciation, and other evolutionary phenomena that could lead to non-Markovian change. Similar arguments could be made about phenotypic evolution. As for the second assumption, correlated evolution among sister species phenotypes can result from ecological interactions (e.g., character displacement, arms races), correlated environments, vicariance leading to convergent evolution, and other factors. These factors are unlikely to be related to phylogenetic relationships among species unless they result from a tendency of related species to show similar evolutionary responses (Harvey and Pagel 1991), for example due to similarities in their genetic or environmental background that bias them towards similar responses (Grafen 1989). Such similarities due to shared background factors are related to inheritance from common ancestors, and can be incorporated explicitly into eqn. (1). If these are not included among the properties of the ancestor in eqn. (1), their effects will appear as correlated evolution among the descendants. Hence the as-

sumption that descendant lineages evolve independently imply that all relevant traits are included in the description of the ancestor.

Our models clearly illustrate how stabilizing selection imposes a phase-forgetting property on the evolutionary process such that historical effects are lost much more quickly than under the other models considered. Only characters evolving under extremely weak stabilizing selection or stabilizing selection with very strong genetic constraints are expected to retain any correlation between phylogenetically related species phenotypes through evolutionary time. If univariate stabilizing selection is the dominant evolutionary force, any remaining relationships among interspecific phenotypes are probably due to other direct responses to change in the environment. If environmental change and changes in the adaptive regime of a group of species are not associated with the phylogeny (as in Lynch and Lande's [1993] model), interspecific measurements may retain little historic information, and thus not be particularly useful in phylogenetic reconstruction. Such data will also not suffer from the usual problems of statistical nonindependence expected in comparative data (e.g., Felsenstein 1985), and can be analyzed quite adequately without reference to any known phylogeny.

As discussed above, not all neutral models of evolution lead to linear or clock-like decreases in phenotypic similarity between species with time. Brownian motion, a reasonable null model of evolution under random genetic drift, leads to a clock-like linear decay of phylogenetic history. However, alternative microevolutionary models formed by modeling the mutation process in slightly different ways (e.g., Cockerham and Tachida 1987) can result in very different macroevolutionary correlation structures. Thus, neutral evolution may not always lead to a linear relationship between phenotypic similarity and time. On the other hand, several different microevolutionary models incorporating selection (e.g., directional, fluctuating directional, and punctuated models) can also produce linear correlation structures and clock-like historic decays. Thus, phylogenetic or comparative methods which assume a linear or clocklike relationship between phenotypic similarity and time need not necessarily assume that the measured characters have been evolving by neutral, gradual change.

The models of directional selection we consider above yield qualitatively the same linear correlation structure as derived from models of neutral evolution. Thus, uniform directional selection, whether constant or fluctuating, may not have any discernible effects on the correlation structure. Without fossils or other historical evidence, it may be impossible to detect a uniform pattern of directional selection from a background of evolution by random genetic drift using interspecific measurements alone.

Phylogenetic methods based on explicit microevolutionary models are preferable to methods based on statistical or algorithmic arguments in that inferring the details of the evolutionary process becomes possible. Our framework provides a first step toward obtaining estimators for the parameters of any microevolutionary model from interspecific measurements. For example, Felsenstein's (1985) method of independent contrasts was proposed as a means of incorporating phylogenies into the statistical analysis of comparative data.

Because his method uses an explicit Brownian motion model of phenotypic evolution, the results have a direct microevolutionary interpretation (Felsenstein 1985, 1988b). Similarly, Martins (1994) developed an estimator for the rate of phenotypic evolution based on the same Brownian motion model. Using our framework and the resulting correlation structures, similar methods can be developed based on alternative models of microevolutionary change. For example, we are presently developing methods to estimate parameters in models of stabilizing selection, such as the strength of the selection and the placement of optima in different environments, using comparative data and the covariance structures described in this paper.

Even the simple models analyzed in this paper provide substantially more flexibility of assumptions than is currently available in most comparative methods. For example, the use of Brownian motion with phylogenetic comparative methods (e.g., Felsenstein 1985) has been criticized because the characters analyzed in comparative studies have usually been subjected to correlated or stabilizing selection along a phylogeny (e.g., Harvey and Pagel 1991). Modern comparative methods that incorporate phylogenetic information into the analysis of interspecific data (e.g., Felsenstein's [1985] independent contrasts and spatial autocorrelation methods [Cheverud et al. 1985; Gittleman and Kot 1991]) are based on specific assumptions about the covariance structure of the comparative data. In theory, these methods can be modified to incorporate any known pattern of covariances (e.g., using generalized least squares, Grafen 1989; Martins and Hansen, In press). If the comparative data are thought to have been generated by a particular microevolutionary process, the approach presented in this paper can be used to derive the covariance structure, which can then be incorporated into an explicit statistical method for comparative analysis. In particular, we plan to investigate the use of covariance structures derived from stabilizing selection or other forms of evolutionary constraints in comparative methods.

Similarly, the evolution of many characters used to reconstruct phylogenetic relationships are likely to have been more constrained than would be expected under a Brownian motion model of evolution. If phenotypic distance is used to infer phylogeny, it may be more appropriate to use exponential rather than linear mappings from phenotype to phylogeny.

Thus, our framework provides (1) a common currency for comparing phylogenetic comparative methods and phylogeny reconstruction algorithms analytically, something that has been done thus far only using computer simulation techniques; (2) a means of uncovering any hidden assumptions underlying phylogenetic methods and linking them with microevolutionary models; and (3) a relatively simple mathematical method that can be used to develop new phylogenetic methods based on alternative views of the microevolutionary process (e.g., methods for reconstructing phylogenies or transforming comparative data based on directional or stabilizing selection; Hansen and Martins, unpubl.). The framework can be used to describe the correlation structure of continuous or categorical characters under a wide variety of microevolutionary processes, and makes only a few minimal assumptions.

LITERATURE CITED

BROOKS, D. R., AND D. H. McLENNAN. 1991. Phylogeny, ecology, and behavior: A research program in comparative biology. Univ. of Chicago Press, Chicago.

CHEETHAM, A. H., J. B. C. JACKSON, AND L.-A. C. HAYEK. 1993. Quantitative genetics of Bryozoan phenotypic evolution. I. Rate tests for random change versus selection in differentiation of living species. Evolution 47:1526–1538.

———. 1994. Quantitative genetics of Bryozoan phenotypic evolution. II. Analysis of selection and random change in fossil species using reconstructed genetic parameters. Evolution 48: 360–375.

CHEVERUD, J. M., M. M. DOW, AND W. LEUTENEGGER. 1985. The quantitative assessment of phylogenetic constraints in comparative analyses: Sexual dimorphism in body weight among primates. Evolution 39:1335–1351.

COCKERHAM, C. C., AND H. TACHIDA. 1987. Evolution and maintenance of quantitative genetic variation by mutations. Proc. Nat. Acad. Sci. USA 84:6205–6209.

EDWARDS, A. F. W. AND L. L. CAVALLI-SFORZA. 1964. Reconstruction of evolutionary trees. Pp. 67–76 in W. H. Heywood and J. McNeill, eds. Phenetic and phylogenetic classification. Syst. Assoc. Publ. No. 6, London.

ELDREDGE, N., AND S. J. GOULD. 1972. Punctuated equilibria: An alternative to phyletic gradualism. Pp. 82–115 in T. J. M. Schopf, ed. Models in paleobiology. Freeman, Cooper and Co., San Francisco, CA.

FELSENSTEIN, J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. Am. J. Human Genet. 25:471–492.

———. 1981. Evolutionary trees from gene frequencies and quantitative characters: Finding maximum likelihood estimates. Evolution 35:1229–1242.

———. 1982. Numerical methods for inferring evolutionary trees. Quar. Rev. Biol. 57:379–404.

———. 1983. Parsimony in systematics: Biological and statistical issues. Annu. Rev. Ecol. Syst. 14:313–333.

———. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

———. 1988a. Phylogenies from molecular sequences: Inferences and reliability. Annu. Rev. Genet. 22:521–565.

———. 1988b. Phylogenies and quantitative characters. Annu. Rev. Ecol. Syst. 19:445–471.

GARDINER, C. W. 1985. Handbook of stochastic methods. 2d ed. Springer Verlag Press, Berlin.

GILLESPIE, J. H. 1991. The causes of molecular evolution. Oxford Univ. Press, Oxford.

GITTLEMAN, J. L., AND M. KOT. 1990. Adaptation: Statistics and a null model for estimating phylogenetic effects. Syst. Zool. 39: 227–241.

GOULD, S. J., AND N. ELDREDGE. 1993. Punctuated equilibrium comes of age. Nature 366:223–227.

GRAFEN, A. 1989. The phylogenetic regression. Phil. Trans. Roy. Soc. Lond. B 326:119–157.

HARVEY, P. H., AND M. D. PAGEL. 1991. The comparative method in evolutionary biology. Oxford Univ. Press, Oxford.

KARLIN, S., AND H. M. TAYLOR. 1975. A first course in stochastic processes. Academic Press, New York.

LANDE, R. 1976a. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

———. 1976b. The maintenance of genetic variability by mutation in a polygenic character with liked loci. Genet. Res. 26:221–235.

———. 1977. Statistical tests for natural selection on quantitative characters. Evolution 31:442–444.

———. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. Evolution 33:402–416.

———. 1985. Expected time for random genetic drift of a population between stable phenotypic states. Proc. Nat. Acad. Sci. USA 82:7641–7645.

———. 1986. The dynamics of peak shifts and the patterns of morphological evolution. Paleobiology 12:343–354.

LOSOS, J. B., AND F. R. ADLER. In press. Stumped by trees? A generalized null model for patterns of organismal diversity. Am. Nat.

LYNCH, M. 1988. The rate of polygenic mutation. Genet. Res. Cambridge 51:137–148.

———. 1989. Phylogenetic hypotheses under the assumption of neutral quantitative-genetic variation. Evolution 43:1–17.

———. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. Am. Nat. 136: 727–741.

———. 1993. Neutral models of phenotypic evolution. Pp. 86–108 in L. Real, ed. Ecological genetics. Princeton Univ. Press, Princeton, NJ.

LYNCH, M., AND W. G. HILL. 1986. Phenotypic evolution by neutral mutation. Evolution 40:915–935.

LYNCH, M., AND R. LANDE. 1993. Evolution and extinction in response to environmental change. Pp. 234–250 in P. M. Karieva, J. G. Kingsolver, and R. B. Huey, eds. Biotic interactions and global change. Sinauer, Sunderland, MA.

MADDISON, D. R. 1994. Phylogenetic methods for inferring the evolutionary history and processes of change in discretely valued characters. Annu. Rev. Entomol. 39:267–292.

MADDISON, W. P., AND D. R. MADDISON. 1992. MacClade: Analysis of phylogeny and character evolution. Sinauer, Sunderland, MA.

MARTINS, E. P. 1994. Estimating rates of character change from comparative data. Am. Nat. 144:193–209.

———. 1995. Phylogenies and comparative data, a microevolutionary perspective. Phil. Trans. Roy. Soc. Lond. B 349:85–91.

MARTINS, E. P., AND T. F. HANSEN. 1996. Phylogenetic comparative methods: The statistical analysis of interspecific data. Pp. 22–75 in E. P. Martins, ed. Phylogenies and the comparative method in animal behavior. Oxford Univ. Press, Oxford.

———. In press. A microevolutionary link between phylogenies and comparative data. In P. Harvey, J. Maynard Smith, and A. Leigh-Brown, eds. New uses for new phylogenies. Oxford Univ. Press, Oxford.

MILES, D. B., AND A. E. DUNHAM. 1993. Historical perspectives in ecology and evolutionary biology: The use of phylogenetic comparative analysis. Annu. Rev. Ecol. Syst. 24:587–619.

SPICER, G. S. 1993. Morphological evolution of the Drosophila virilis species group as assessed by rate tests for natural selection on quantitative characters. Evolution 37:1240–1254.

SWOFFORD, D. L., AND G. J. OLSEN. 1990. Phylogeny reconstruction. Pp. 411–501 in D. M. Hillis and C. Moritz, eds. Molecular systematics. Sinauer, Sunderland, MA.

TURELLI, M., J. H. GILLESPIE, AND R. LANDE. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. Evolution 42:1085–1089.

WILEY, E. O. 1981. Phylogenetics: The theory and practice of phylogenetic systematics. John Wiley and Sons, New York.

Corresponding Editor: D. Maddison

APPENDIX A

The Interspecific Covariance Equation

If phenotypic evolution along a phylogeny is modeled as a Markovian stochastic process unfolding along a tree-like branching structure, with processes continuing along each of the daughter branches resulting from a fork in the tree, with the value at that fork as initial value, then:

$$\text{Cov}[\mathbf{X}_i, \mathbf{X}_j] = \text{Cov}[E[\mathbf{X}_i|\mathbf{X}_z], E[\mathbf{X}_j|\mathbf{X}_z]]$$
$$+ E[\text{Cov}[\mathbf{X}_i, \mathbf{X}_j|\mathbf{X}_z]] \qquad (A1)$$

where $\mathbf{X}_i$, a column vector of traits, is the state of species i, and z denotes the most recent common ancestor of species i and j.

Proof: Using a formula analogous to that for the expectation of conditioned variances, we may write:

$$E[\text{Cov}[\mathbf{X}_i, \mathbf{X}_j | \mathbf{X}_z]] = E[E[\mathbf{X}_i\mathbf{X}_j | \mathbf{X}_z]] - E[E[\mathbf{X}_i | \mathbf{X}_z]E[\mathbf{X}_j | \mathbf{X}_z]]$$

$$= E[\mathbf{X}_i\mathbf{X}_j] - E[E[\mathbf{X}_i | \mathbf{X}_z]E[\mathbf{X}_j | \mathbf{X}_z]] \pm E[\mathbf{X}_i]E[\mathbf{X}_j]$$

$$= \text{Cov}[\mathbf{X}_i, \mathbf{X}_j] - \text{Cov}[E[\mathbf{X}_i | \mathbf{X}_z], E[\mathbf{X}_j | \mathbf{X}_z]]. \qquad (A2)$$

Remark: If the evolutionary changes leading to the state of species i $(\mathbf{X}_i)$ is independent of the evolutionary changes leading to the state of species j $(\mathbf{X}_j)$ after their most recent common ancestor, z, $\text{Cov}[\mathbf{X}_i, \mathbf{X}_j|\mathbf{X}_z]$ = 0. Thus, the second term in (A1) vanishes. The result can also be extended to non-Markovian processes by conditioning, not just on the most recent common ancestor, but on as many common ancestors as necessary. The only change to eqn. (A1) is that $\mathbf{X}_z$ must then be replaced by the collection of these common ancestors.

APPENDIX B
Gaussian Processes

A diffusion process where the infinitesimal mean and variance are linear functions of the state variables is called linear. With suitable initial conditions, any collection of points resulting from a linear diffusion process will have a multivariate normal distribution and the process is termed "Gaussian." If a Gaussian process unfolds on a tree, it can be shown (Appendix C) that any collection of points on the tree is multivariate normal so that the collection of points is completely characterized by their mean vector and variance-covariance matrix. Gaussian processes yield linear regressions between points on the phylogeny (such as those in eqn. [1]) which lead to comparably simple formulae for the covariances between trait values measured in different species.

Consider a linear diffusion with an infinitesimal mean vector $\mu(\mathbf{X})$ = $\mathbf{a} - \mathbf{AX}$ and a constant infinitesimal variance-covariance $\sigma^2(\mathbf{X})$ = $\mathbf{B}$; this is general enough to contain all examples in the main text. Using standard techniques (e.g., Gardiner 1985) the mean vector at time t, $\mathbf{m}(t) = E[\mathbf{X}(t)]$, is given as the unique solution of the initial value problem $d\mathbf{m}/dt = \mathbf{a} - \mathbf{Am}$, $\mathbf{m}(0) = E[\mathbf{X}(0)] = \mathbf{m}_0$. This solution is $\mathbf{m}(t)$ = $\mathbf{Q}(t)\mathbf{m}_0 + (\mathbf{I} - \mathbf{Q}'(t))\mathbf{A}^{-1}\mathbf{a}$, where $\mathbf{Q}(t) = \text{Exp}[-\mathbf{A}t]$, and $\mathbf{Q}'(t)$ = $\text{Exp}[-\mathbf{A}'t]$. This leads directly to the linear regression $E[\mathbf{X}(t + s)|\mathbf{X}(s)]$ = $\mathbf{Q}(t)\mathbf{X}(s)$. Using eqn. (1), we find that:

$$\text{Cov}[\mathbf{X}_i, \mathbf{X}_j] = \mathbf{Q}(t_{iz})\mathbf{V}(t_z)\,\mathbf{Q}'(t_{jz}), \qquad (B1)$$

where $\mathbf{V}(t_z) = \text{Var}[\mathbf{X}_z]$ is the variance-covariance matrix of traits in taxon z, the most recent common ancestor of species i and j. If the $\mathbf{A}$ matrix contained in $\mathbf{Q}$ has only eigenvalues with positive real part, the covariances between species show an exponential decay with time such that species that are further in time from their common ancestor will be less similar to one another than will closely related species.

The matrix $\mathbf{V}(t)$ is given as the solution of the initial value problem $d\mathbf{V}(t)/dt = -\mathbf{AV}(t) - \mathbf{V}(t)\mathbf{A}' + \mathbf{B}$, $\mathbf{V}(0) = \mathbf{V}_0$, where $\mathbf{V}_0$ is the initial condition. Note that this is independent of the $\mathbf{a}$ vector. Using the zero matrix as an initial value assumes that the process starts from a fixed value at the root of the tree. We may also assume that it starts from a normal initial distribution, in which case $\mathbf{V}_0$ should be taken as the variance-covariance matrix of this distribution. The solution to this matrix initial value problem is:

$$\mathbf{V}(t) = \int_0^t \mathbf{Q}(t - \tau)\mathbf{B}\mathbf{Q}'(t - \tau)\,d\tau + \mathbf{Q}(t)\mathbf{V}_0\mathbf{Q}'(t). \qquad (B2)$$

The process admits a stationary solution when all eigenvalues of $\mathbf{A}$ have positive real part. In this case (B2) simplifies to

$$\mathbf{V}(t) = \hat{\mathbf{V}} - \mathbf{Q}(t)\,(\hat{\mathbf{V}} - \mathbf{V}_0)\mathbf{Q}'(t), \qquad (B3)$$

where $\hat{\mathbf{V}}$ is the equilibrium variance-covariance matrix of the process. It is given implicitly by the matrix equation $\mathbf{A}\hat{\mathbf{V}} + \hat{\mathbf{V}}\mathbf{A}' = \mathbf{B}$. If $\mathbf{A}^{-1}\mathbf{B}$ is symmetric, as is the case in our applications, the solution to this equation is $\hat{\mathbf{V}} = \mathbf{A}^{-1}\mathbf{B}/2$. In general there is no solution for $\hat{\mathbf{V}}$ in terms of $\mathbf{A}$ and $\mathbf{B}$, and $\hat{\mathbf{V}}$ must then be obtained by solving the equivalent linear system of equations for its elements. If $\mathbf{A} = \mathbf{0}$, as in a Brownian motion, no stationary solution exist, and the solution to (B2) is simply $\mathbf{V}(t) = \mathbf{V}_0 + \mathbf{B}t$.

Using (B3) in (B1), and assuming that the variance at the root of the tree is zero $(\mathbf{V}_0 = \mathbf{0})$ gives:

$$\text{Cov}[\mathbf{X}_i, \mathbf{X}_j] = \mathbf{Q}(t_{iz})\hat{\mathbf{V}}\mathbf{Q}'(t_{jz}) - \mathbf{Q}(t)\hat{\mathbf{V}}\mathbf{Q}'(t), \qquad (B4)$$

where t is the time from the root of the tree to present. Note that the last term in this equation is the same for all species and equals zero if t is large, i.e., if the process has reached equlibrium. Equation (B4) specifies a matrix in which the entries are the covariances of all trait combinations of two species. For example, the covariance between trait k in species i and trait l in species j is:

$$\text{Cov}[X_{ik}, X_{jl}] = [\mathbf{Q}(t_{iz})]_k\,\hat{\mathbf{V}}[\mathbf{Q}(t_{jz})]'_l - [\mathbf{Q}(t)]_k\,\hat{\mathbf{V}}[\mathbf{Q}(t)]'_l, \qquad (B5)$$

where the notation $[\,]_i$ means the ith row of matrix and the transpose is of this row vector and not the matrix.

APPENDIX C
Phylogenetic Distribution under Gaussian Processes

If evolution unfolds on a tree according to a Markov Gaussian process and branches evolve independently after their split, then the joint distribution of any set of taxon phenotypes on the tree is multivariate normal (MVN).

Proof: Let $\mathbf{X} = \{\mathbf{X}_1, .., \mathbf{X}_n\}$ be the vector of species values for n species. Each species value can be multivariate such that $\mathbf{X}_i = \{X_{i1}, .., X_{im}\}$ is the vector of m trait values of species i. The aim is to show that the distribution of $\mathbf{X}$, denoted $p(\mathbf{X})$, is MVN. We do this inductively by showing that if one has a collection of species like $\mathbf{X}$ that is MVN and includes the root species, and add one more species, $\mathbf{Y}$, that is not ancestor to any of the species in $\mathbf{X}$, then the joint distribution of $\mathbf{X}$ and $\mathbf{Y}$ is MVN. As any collection of species can be constructed in this way by successively adding one and one species that is not ancestor to any of the previous species, and as the root species has a Gaussian distribution, the desired result follows by induction. If the root is not in $\mathbf{X}$, the result follows as marginals of normals are normals.

The induction step is established as follows. The joint density of $\mathbf{X}$ and $\mathbf{Y}$ is $p(\mathbf{X}, \mathbf{Y}) = p(\mathbf{Y}|\mathbf{X})p(\mathbf{X})$. We know by assumption that $p(\mathbf{X})$ is MVN. By the Markov property, and the independence of branches after their split, $p(\mathbf{Y}|\mathbf{X}) = p(\mathbf{Y}|\mathbf{X}_z)$, where $\mathbf{X}_z$ is the most recent common ancestor of $\mathbf{Y}$ among $\mathbf{X}$. This common ancestor always exists, as the root is included in $\mathbf{X}$. The function $p(\mathbf{Y}|\mathbf{X}_z)$ is the transition density under a Gaussian process and hence MVN in $\mathbf{Y}$. The exponent of this distribution is $-1/2(\mathbf{Y} - \mathbf{V}_{yx}\mathbf{V}^{-1}_x\mathbf{X}_z)(\mathbf{V}_y - \mathbf{V}_{yx}\mathbf{V}^{-1}_x\mathbf{V}'_{yx})^{-1}(\mathbf{Y} - \mathbf{V}_{yx}\mathbf{V}^{-1}_x\mathbf{X}_z)'$, where $\mathbf{V}_x$, $\mathbf{V}_y$ and $\mathbf{V}_{yx}$ are variance-covariance matrices of $\mathbf{X}_z$, $\mathbf{Y}$, and $\mathbf{Y}$ with $\mathbf{X}_z$, respectively. We have for simplicity assumed that the marginal means are zero. This exponent is a quadratic form in $\mathbf{X}$ and $\mathbf{Y}$ and as multiplication with $p(\mathbf{X})$ amounts to adding the exponents we observe that the exponent of $p(\mathbf{X}, \mathbf{Y})$ is a quadratic form. This quadratic form is negative definite, as its coefficient matrix must be the inverse of the variance-covariance matrix of $\{\mathbf{Y}, \mathbf{X}\}$. Thus $p(\mathbf{X}, \mathbf{Y})$ is MVN.

APPENDIX D
Renewal Counting Processes

A renewal counting process (e.g., Karlin and Taylor 1975) counts the number of events, N(t), that have happened during a time interval, (0, t). The essential assumption is that the interevent times are independent and identically distributed. We need, as a slight extension, to consider a "delayed" renewal process where the first interevent time is allowed to have a different distribution from the rest. The overall mean of the process, $E[N(t)]$, is called the "renewal function." The regression on an ancestor for this process is $E[N_i(t_i)|N_z(t_z)] = N_z(t_z) + E[N(t_{zi})]$. The covariance between the number of evolutionary events leading to the phenotypes of two species can then be computed with eqn. (1):

$$\text{Cov}[N_i, N_j] \approx \text{Cov}[N_z + E[N(t_{zi})], N_z + E[N(t_{zj})]]$$
$$= \text{Cov}[N_z, N_z] = \text{Var}[N_z], \qquad (D1)$$

where N_i is the state of the process in species i, and z is the most recent common ancestor of species i and j. Equation (D1) is exact only when the counting process is a Poisson process, otherwise it is an approxi-

mation because knowledge of states before the ancestor z may carry some information about the distribution of the first interevent time after t_z, causing the process to violate the Markovian assumption. However, if the number of evolutionary changes occurring along each branch of the phylogeny is not very small, eqn. (D1) should be a very good approximation. Hence, for renewal counting processes, the covariance between species equals the variance of their most recent common ancestor.

If the renewal process is "stationary" (i.e., if it began indefinitely far in the past, or we began the counting at a random point in time), then N(t) is normally distributed with mean and variance proportional to time such that:

$$E[N(t)] \sim \frac{t}{\mu}, \qquad Var[N(t)] \sim \frac{\sigma^2}{\mu^3} t, \qquad (D2)$$

where μ and σ^2 are the mean and variance of the interevent times (Karlin and Taylor 1975). Hence,

$$Cov[N_i, N_j] \approx \frac{\sigma^2}{\mu^3} t_z = \frac{\sigma^2}{\mu^3}\left(t - \frac{t_{ij}}{2}\right) \qquad (D3)$$

and $Corr[N_i, N_j] \approx 1 - t_{ij}/2t$. Thus the covariance between species in the number of evolutionary changes that have occurred decays linearly with distance between species.

The species mean phenotypes are sums of the evolutionary changes that have occurred through time, and are given by

$$X(t) = \sum_{k=1}^{N(t)} Y_k + X(0), \qquad (D4)$$

where N(t) is the counting process as above, the Y_k are the set of evolutionary changes in the mean phenotype of the species occurring at each interval (these are assumed to be independent, identically distributed random vectors that are also independent of N([t]), and $X(0)$ is the mean phenotype at the root of the tree. We can obtain the regressions of descendants on ancestors in Eqn. 1 by considering the cumulant generating function (CGF) of $X(t) - X(0)$. It can be shown that $\phi_X(z) = \phi_N(\phi_Y(z))$, where $\phi_N(z)$ and $\phi_Y(z)$ are the CGFs of N and Y, respectively. If we then take the first and second derivative of $\phi_X(z)$ and evaluate at $z = 0$, we find that:

$$E[X(t)] = E[Y]E[N(t)] \text{ and}$$
$$Var[X(t)] = Var[Y]E[N(t)] + E[Y]E[Y]'Var[N(t)], \qquad (D5)$$

where E[Y] is the mean vector and Var[Y] the variance-covariance matrix of the Y vector. Using $E[X(t + s)|X(s)] = E[Y]E[N(t)] + X(s)$ in eqn. (1) yields

$$Cov[X_i, X_j] = Var[X_z]$$
$$= Var[Y]E[N(t_z)] + E[Y]E[Y]'Var[N(t_z)]$$
$$= \frac{1}{\mu}\left(Var[Y] + \frac{\sigma^2}{\mu^2}E[Y]E[Y]'\right)t_z. \qquad (D6)$$