Palaeontologia Electronica
http://palaeo-electronica.org

ON THE SIMULATION OF THE EVOLUTION OF MORPHOLOGICAL SHAPE: MULTIVARIATE SHAPE UNDER SELECTION AND DRIFT

P. David Polly

ABSTRACT

Stochastic computer simulation is an important method for comparing the evolutionary patterns and processes associated with radically different intervals of time. This paper demonstrates how to simulate the evolution of complex morphologies over geological timescales of millions of generations. The simulations are used to test how various assumptions about microevolutionary parameters and processes manifest themselves on macroevolutionary timescales. Complex morphology is modelled using geometric representations of shape (e.g., landmarks or outlines), and so the procedure described here is limited to single rigid structures. The procedure is based on empirically measured phenotypic correlations, which constrain the evolutionary outcomes in biologically realistic ways. Different microevolutionary assumptions about covariances, population size, and evolutionary mode can be tested by incorporating them into the simulation parameters.

The evolution of molar tooth morphology in shrews is simulated under four different evolutionary modes: (1) randomly fluctuating selection; (2) directional selection; (3) stabilizing selection; and (4) genetic drift. Each of these modes leaves a distinctive imprint on the distribution of morphological distances, a feature that can be used to reconstruct the mode from real comparative data. A comparison of the results with real data on shrew molar diversity suggests that teeth have evolved predominantly by randomly fluctuating selection. The rate of divergence in shrew molars is greater than expected under drift, but it is neither linear nor static as expected with directional or stabilizing selection.

The evolution of morphology with randomly fluctuating selection is also simulated on a phylogenetic tree. Daughter species share derived morphologies and positions within the principal components spaces in which the simulation is run. This result suggests that phylogeny can be successfully reconstructed from multivariate morphometric data when organisms have evolved under any mode except strong stabilizing selection.

School of Biological Sciences, Queen Mary, University of London,
London E1 4NS, U.K. d.polly@qmul.ac.uk,

KEY WORDS: evolutionary rates, genetic drift, molar teeth, morphological evolution, phenotypic covariances, selection, *Sorex* (Soricidae, Lipotyphla, Mammalia), stasis, shrew

Whitey Hagadorn was sole Executive Editor for this article.

Polly, P. David, 2004. On the Simulation of the Evolution of Morphological Shape: Multivariate Shape Under Selection and Drift, *Palaeontologia Electronica* Vol. 7, Issue 2; 7A:28p, 2.3MB;
http://palaeo-electronica.org/paleo/2004_2/evo/issue2_04.htm

PE Article Number: 7.2.7A

Submission: 23 August 2004 Acceptance: 27 November 2004

INTRODUCTION

Stochastic computer simulation is an important method for comparing the evolutionary patterns and processes associated with radically different intervals of time (Raup and Gould 1974; Ibrahim et al. 1995). Simulations allow the extrapolation of microevolutionary processes to macroevolutionary timescales: populations can be evolved for millions of generations using particular parameters (effective population sizes, phenotypic variances, and heritabilities) and processes (directions and intensities of selection, drift). The adequacy of population-level phenomena for explaining macroevolutionary patterns can be tested by comparing the results with real morphological diversity, or, conversely, population-level parameters can be reverse engineered from comparative data by systematically varying the parameters used in the simulations. Whereas simulations have been devised for comparatively simple traits—such as size, shell shape, or gene frequencies—a procedure for modelling the long-term evolution of complicated morphologies using quantitative genetic parameters remains difficult for most palaeontological researchers.

This paper demonstrates how to perform a simulation of the evolution of almost any rigid morphological shape that can be represented geometrically (e.g., using landmarks or outlines in either two or three dimensions), regardless of its complexity. A crucial requirement is the availability of a sample from which population-level phenotypic covariances can be estimated. Directional, stabilizing, and random selection are simulated as walks in the empirical variation space defined by the principal components of the phenotypic covariance matrix. The morphological shape, or phenotype, at any step of the simulation can be retrieved, visualized, or utilized for quantitative comparisons. Simulations can be run multiple times for millions of generations to demonstrate the effects of microevolutionary processes over palaeontological timescales. The simulated morphologies can be used for statistical comparison with real comparative data from the fossil (or extant) record.

The simulation procedure is demonstrated by looking at the effects of directional selection, stabilizing selection, randomly fluctuating selection, and genetic drift on the long-term evolution of mammalian molar shape. Each of these evolutionary modes leaves a distinctive imprint on the distribution of morphological shape, a feature that can be used to estimate the mean mode that produced real morphological shapes in a particular lineage or clade. In this paper, I compare the outcomes of simulations of the evolution of molar crown shape in a single species of shrew, *Sorex araneus* (Soricidae, Lipotyphla, Mammalia) to the real diversity of molar shape in species belonging to the same clade. A simulation of randomly fluctuating selection is also applied to a branching phylogeny in order to determine how derived morphometric shape accrues in independent clades.

Sorex araneus, the Eurasian common shrew, is a widespread and genetically interesting animal. Today, the species is spread over most of Europe and across Siberia and Central Asia beyond Lake Baikal. It is genetically subdivided into more than 70 named chromosomal races (Wójcik et al. 2003), and Quaternary fossil samples can be assigned to broad subgroups within these races with reasonable accuracy (Polly 2003a). *S. araneus* is part of a larger species complex that is peculiar in that males typically have an extra X chromosome, in addition to the normal mammalian complement of XY (Zima et al. 1998). In humans, this condition is abnormal and referred to as Kleinfelter's syndrome. The *S. araneus* complex evolved no more than 2.5 million years ago (Fumagalli et al. 1999). Species within the group – which include *S. coronatus*, *S. granarius*, *S. samniticus*, *S. arcticus*, and *S. tundrensis* – are often difficult to distinguish on non-morphometric morphological grounds; hence, the fossil specimens have been usually referred to *S. araneus sensu lato*. *S. araneus* and its relatives have a rich fossil record covering three continents (Harris 1998; Rzebik-Kowalska 1998; Storch et al. 1998). The earliest probable member of the broader group is from the Late Pliocene of Schernfeld, Germany (Dehm 1962), and by the mid-Pleistocene they were widespread in Europe and Asia. In North America the group was present in the form of *S. arcticus* by the Late Pleistocene as far south as Arkansas and east as West Virginia (Kurtén and Anderson 1980; Semken 1984; Harris 1998). Typically the species reproduces at a rate of one or two generations per year.

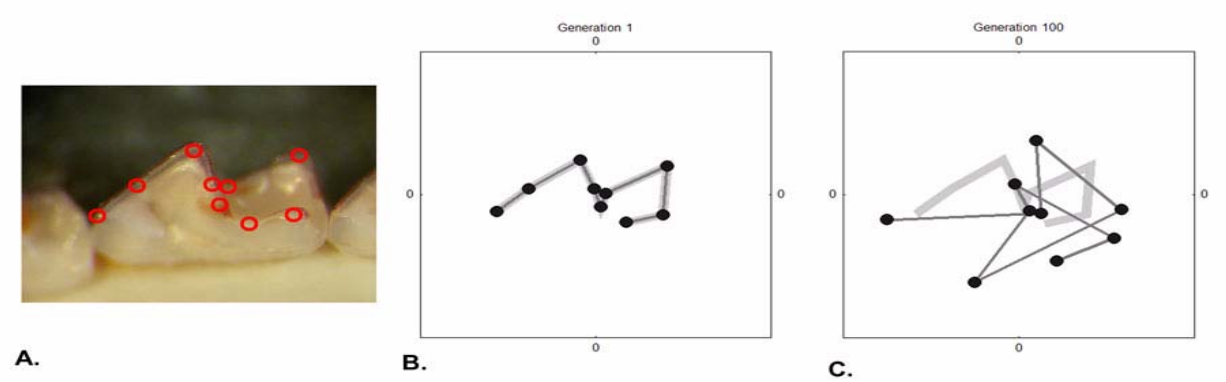

Figure 1. The effect of simulating the evolution of morphological shape without a covariance matrix. **A.** Landmarks used to represent the shape of the crown of the upper first molar of *Sorex araneus* (Lipotyphla, Mammalia). **B.** Configuration of landmarks at the beginning of the simulation. Black dots connected by thin, dark lines show the positions of landmarks, and the thick, light grey line represents the initial position. **C.** The configuration of landmarks after 100 generations of random evolution. Anatomically adjacent landmarks become unrealistically entwined when structural correlations are not included in the simulation model. **D.** Animation of the complete simulation.

The Importance of Phenotypic Covariances

Empirically measured phenotypic covariances are an important part of this simulation procedure because they overcome a fundamental difficulty: morphological structures have patterns of correlation that channel their variation and evolution in non-random ways (Olson and Miller 1958; Zelditch et al. 1992). Parts of a biological structure do not vary randomly with respect to one another; rather, their topographical relationships are constrained by growth, genetics, function, materials, and general physical laws (Maynard Smith et al. 1985; Arnold 1992). Simulations that do not take into account these correlations quickly produce biologically unrealistic morphologies (Figure 1).

While it is difficult to link correlations to their various possible causes, it is relatively straightforward to measure their combined effects on a species or population. The phenotypic covariance matrix (**P**) describes the amount of variation (variances) in each part of a complex structure and the correlation (covariance) among the parts. **P** provides information on the correlated response of parts of the structure relative to one another. Traits that are united by a particular constraint will have high covariances, while those that are uncon-

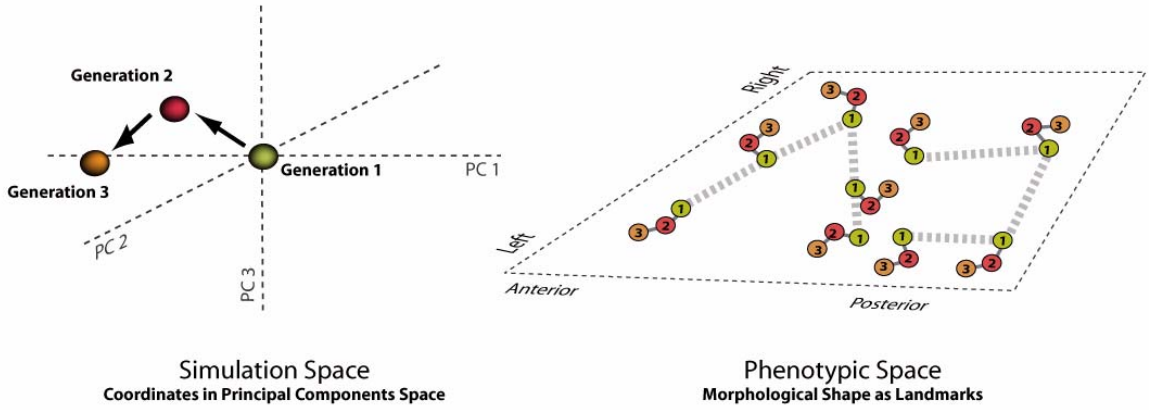

Figure 2. Each simulation is run in a principal component space (left), whose positions correspond to unique configurations of landmarks in the phenotypic space (right). The diagram illustrates three steps in a simulation of a molar tooth with nine landmarks (*c.f.*, Figure 1). As the simulation moves through the PC space at left, the landmarks of the phenotype at right change according to the direction and size of the steps along each of the PC axes. Different phenotypic correlations are associated with each of the several PC axes, and the change in morphology at each step is the sum total of the changes on all axes. Different evolutionary modes (e.g., directional selection) are modeled by biasing the direction and length of the individual steps in the principal component space.

strained will have little or no covariance. **P** can be measured from a sample of individuals randomly drawn from a population or species, such as the individuals found in many museum collections. The number required will depend on how extensive variation is in the particular structure being studied, but for vertebrate traits such as teeth and skulls a sample size of 10 can be minimally sufficient, although 20 to 50 individuals will yield a more accurate estimate.

The simulation procedure described in this paper is simplified by making use of the principal components of **P**, rather than **P** itself. Every covariance matrix can be described as a series of principal components that represent independent components, or portions, of the variation and covariation in the original matrix. In technical terms, principal components are a series of orthogonal (or uncorrelated) axes known as *eigenvectors*, each of which describes a particular amount of the total variance as indicated by the *eigenvalues* (Blackith and Reyment 1971; Tatsuoka and Lohnes 1988). The advantage to the principal components is that a univariate simulation can be run for each of the eigenvectors and the results summed over all components to produce the full multivariate simulation (Figure 2). This shortcut is possible because the individual components are statistically independent of one another, and it greatly reduces the number of calculations that are required at each step of the simulation. When the simulations span hundreds of thousands or millions of steps, the savings are important. While the principal component space of the example is only shown in three dimensions, this procedure may be carried out across as many as applicable to a given dataset. In the simulations presented below, shape variation is evolved in 15 eigenvector dimensions.

The use of the principal components also overcomes a difficulty associated with the fact that phenotypic covariance matrixes of geometric shape are singular. Singular **P** matrices result when some variables do not have any variation of their own which is independent of the others. All geometric morphometric covariance matrixes are singular because some of the biological variation – size, orientation, and position – is removed from the data by superimposition (Gower 1975; Rohlf and Slice 1990; Rohlf 1999). A singular matrix can be described by a set of principal components whose number is smaller than the number of variables in the original **P** matrix. This reduction in the number of variables also improves the computational efficiency of the simulation. Principal components of a singular matrix can be found using a standard procedure called Singular Value Decomposition (Golub and Van Loan 1983).

Relation to Models of Multivariate Evolution

Lande (1976, 1979) pioneered the study of the evolution of correlated characters. He provided equations for the response of phenotypes to selection and drift that were based on the concepts of adaptive landscapes first proposed by Simpson

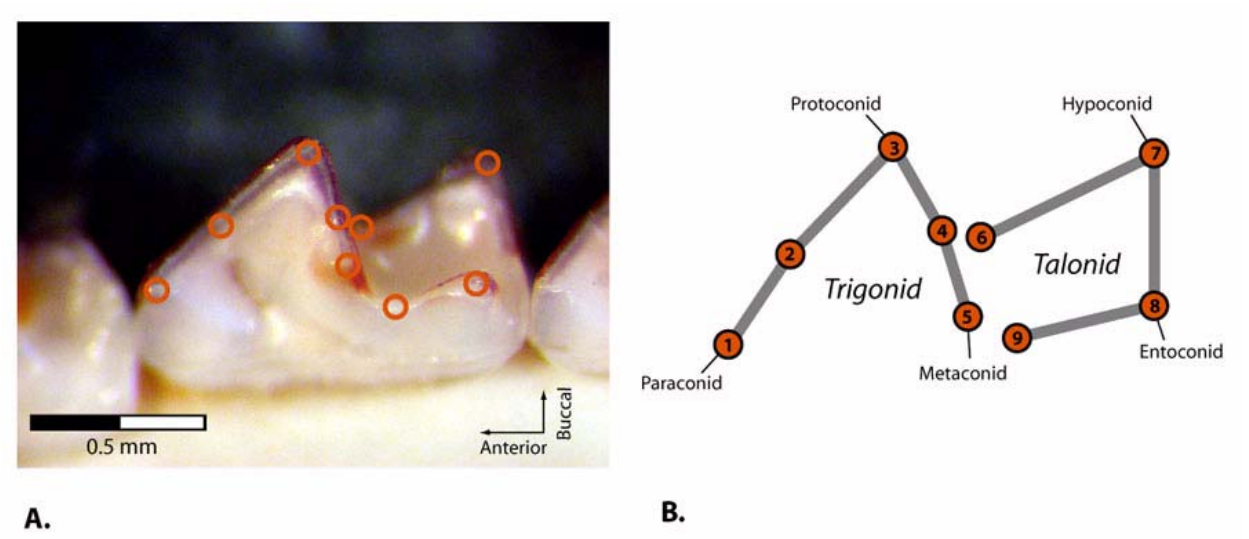

Figure 3. Landmarks used to represent tooth crown morphology. **A.** Crown of the first lower right molar of *Sorex araneus* in functional occlusal view, in which the tooth is oriented so that the line of sight is parallel to all of the vertical shearing blades. The positions of the nine landmarks are indicated by open orange circles. **B.** Diagrammatic representation of the nine landmarks as portrayed in many of the simulation results. The trigonid and talonid landmarks are each connected by a thick grey line. The five cusp landmarks are labelled.

(1944, 1953) and Wright (1968). He showed that the multivariate response to selection depends on the heritable, or additive genetic, covariances among the traits, as well as on the direction and magnitude of selection applied to each trait. Lande's multivariate approach has been widely used in microevolutionary studies (Lande 1986; Lande and Arnold 1983; Atchley and Hall 1991; Cheverud 1995; Marroig and Cheverud 2001; Klingenberg and Leamy 2001; Arnold et al. 2001), but seldom used to explore evolution over timescales as large as those in the fossil record (but see Cheetham et al. 1993, 1994, 1995).

The simulation method described here is a modification of Lande's evolutionary model. Lande (1979) described the evolution of multivariate morphology as

$$\Delta\bar{z} = \beta G \quad (1)$$

where $\Delta\bar{z}$ is the change in the mean morphology of a population over the space of a single generation, **G** is the additive genetic covariance matrix, and β is the series of selection coefficients that are applied to each variable used to describe the morphological structure. **G** represents that portion of **P** which is heritable, or passed, from parents to their offspring. β is a set of selection coefficients (or drift coefficients), one for each of the variables used to describe the phenotype. If **G** is known and a particular set of β is applied to it, then the change in the mean phenotype for that generation is $\Delta\bar{z}$.

Unfortunately, **G** is unknown for most morphological structures and for all palaeontological species. However, we can substitute **P** for **G** by adding an extra parameter, heritability, to the simulation. The basic formula for morphological change at each generation in the simulations is, therefore,

$$\Delta\bar{z} = \beta HP \quad (2)$$

where β is the vector of selection coefficients, and **H** is a heritability matrix that transforms **P** into **G**. Because **H** is also unknown, it must be simulated along with β. It is convenient to simulate β**H** as a single vector (β**H** is a vector the same length as β), rather than as both a vector and a matrix because it reduces the number of computations and because the elements of vector β**H** are equivalent to per-generation rates of morphological change in standardized variance units, which are commonly reported in palaeontological or comparative studies (Gingerich 1993; Kingsolver et al. 2001).

This simulation thus does not represent a "constant heritability" model (Lande 1976; Spicer 1993) because **H** forms part of the random parameter. It is also not a mutation-drift-equilibrium model (Turelli et al. 1988; Spicer 1993) because the phenotypic variance in each generation is held constant as part of **P**. In nature, variance in a population is a balance between that which is removed by selection and that which is produced by muta-

tion and recombination. In principle, a mutation-drift model could be simulated, but the parameter of the mutational variance would be difficult to estimate for shrew molars. In practice, the assumption of a stochastically constant variance is not unrealistic because the total variance in shrew molar shape does not change dramatically over long periods of evolutionary time (Polly 2005).

Technical discussions of the effect of transforming **P** into its principal components and the effects of simulations on a singular **P** matrix are found in the Appendix.

MATERIALS AND METHODS

The starting shape and **P** matrix for the simulations were estimated from a sample of the European Common shrew, *Sorex araneus*, from Bialowieza National Forest, Poland (N = 43), housed in the Mammal Research Institute of the Polish Academy of Sciences at Bialowieza. The first molar of each individual was photographed through a microscope in functional occlusal view (Butler 1961; Crompton 1971), a position that is less prone to orientation error than other views (Polly 2001a, 2003a). Each specimen was photographed five times and averaged to minimize further the error due to orientation and landmark placement. Nine two-dimensional landmarks were placed on major crown cusps and notches to represent the morphology of the tooth (Figure 3) and superimposed using Procrustes (GLS) superimposition (Gower 1975; Rohlf 1999). In principle, this simulation could be carried out with 3-D representations of shape, but two dimensions were used here to facilitate comparison with existing data on variation in shrew molar morphology. The mean of the superimposed teeth was used as the starting morphology for each simulation.

P was estimated as the covariance matrix of the Procrustes residuals of the superimposed landmarks that remain after the mean shape is subtracted (Table 1). With nine two-dimensional landmarks, there were a total of 18 residual shape variables, making **P** an 18 x 18 matrix. **P** only has 15 non-zero eigenvectors, however, because three dimensions are lost due to the removal of variation in size, translation, and rotation during superimposition (Table 2). Each simulation, therefore, required a vector β**H** with 15 elements, one for each dimension of the PC space.

The elements of β**H** were randomly generated at each step of the simulation. Each simulation drew elements from a distribution appropriate to the mode of evolution being modelled, as described below. The median β**H** for all of the simulations except genetic drift was based on the rate of evolution of molar shape in *Cantius* (Primates, Mammalia) from the early Eocene of Wyoming (Clyde and Gingerich 1994). These authors estimated the average per-generation rate of change in Bookstein shape coordinates using the log-rate/log-interval (LRI) method (Gingerich 1993). LRI rates are reported in units of phenotypic standard deviations per generation, or *haldanes*, which are equivalent to the square root of β**H**. Gingerich and Clyde estimated a rate of 0.653 *haldanes* per landmark dimension, the squared value of which (0.426) was used as the standard deviation for most of the β**H** distributions in this paper. The accuracy of this rate and its appropriateness for shrew molars are arguable, but it is the only empirical estimate of the rate of evolutionary change in molar shape that is available.

Effective population size, N_e, is the main parameter required for simulating genetic drift (Wright 1931). Drift occurs when population means differ from one generation to the next because some individuals fail to reproduce by chance: the smaller the population size, the greater the effect. The expected change due to drift is equal to the heritable morphological variance divided by N_e (Lande 1976). N_e is not easy to estimate in living animals because population sizes fluctuate over time. In practice, N_e is estimated either from population censuses or from molecular markers (Avise et al. 1988; Avise 1994, 2000); the latter usually produces much larger estimates because it represents an average over large geographic and temporal scales. The simulation of drift was run twice, once with N_e = 70, an estimate based on population censuses (Churchfield 1982, 1990, personal commun., 2004; Wójcik, personal commun., 2004), and once with N_e = 70,000, an estimate based on molecular estimates (Ratkiewiecz et al. 2002).

Each simulation required the random β**H** elements to be generated in a different way. For the simulation of random selection, the elements were drawn from a normal distribution centred on zero with a standard deviation of 0.426. This distribution yields numbers that are equally likely to be positive or negative, with a mean absolute magnitude of 0.338. Directional coefficients were drawn from a skewed distribution, with the probability of a negative value marginally higher than a positive one, but still with an absolute magnitude of 0.338. A negative bias was arbitrarily applied to all eigenvectors, though directional selection would still obtain if the direction were positive in some dimensions and negative in others. Stabilizing selection

Table 1. Phenotypic covariance matrix (P) based on the Bialowieza sample of *Sorex araneus*. Landmark numbers correspond to Figure 3. Note that variances and covariances are small because they are in Procrustes shape units, where 1.0 equals the size of the entire shape as measured by the sum of squared distances from each landmark to the shape centroid.

	1X	1Y	2X	2Y	3X	3Y	4X	4Y	5X	5Y	6X	6Y
1X	8.8E-05	1.1E-05	-2.0E-05	-2.2E-05	-8.9E-06	4.3E-05	-8.4E-06	1.5E-05	-3.4E-05	1.7E-05	-8.0E-06	-2.4E-05
1Y	1.1E-05	6.3E-05	-1.7E-05	-7.7E-06	-8.6E-06	-2.2E-05	-1.5E-05	-2.8E-05	9.4E-06	-9.2E-06	-1.5E-05	-1.6E-05
2X	-2.0E-05	-1.7E-05	5.0E-05	2.6E-05	-5.2E-06	-1.7E-05	-1.2E-05	9.4E-07	-9.1E-06	-1.3E-05	-1.8E-05	1.8E-05
2Y	-2.2E-05	-7.7E-06	2.6E-05	8.6E-05	-1.2E-05	-6.2E-05	5.8E-06	-1.8E-05	1.8E-05	-8.1E-05	-1.4E-05	2.8E-05
3X	-8.9E-06	-8.6E-06	-5.2E-06	-1.2E-05	5.8E-05	1.4E-07	1.3E-05	-1.7E-05	7.2E-07	1.5E-06	-1.6E-05	-1.5E-05
3Y	4.3E-05	-2.2E-05	-1.7E-05	-6.2E-05	1.4E-07	1.6E-04	-3.2E-06	1.8E-05	-1.2E-05	4.3E-05	2.2E-05	-6.2E-05
4X	-8.4E-06	-1.5E-05	-1.2E-05	5.8E-06	1.3E-05	-3.2E-06	2.8E-05	-1.1E-05	1.3E-05	-1.0E-05	4.0E-06	9.0E-06
4Y	1.5E-05	-2.8E-05	9.4E-07	-1.8E-05	-1.7E-05	1.8E-05	-1.1E-05	6.5E-05	-1.7E-05	2.3E-05	2.8E-05	8.0E-06
5X	-3.4E-05	9.4E-06	-9.1E-06	1.8E-05	7.2E-07	-1.2E-05	1.3E-05	-1.7E-05	7.8E-05	-4.4E-05	-1.5E-05	8.9E-06
5Y	1.7E-05	-9.2E-06	-1.3E-05	-8.1E-05	1.5E-06	4.3E-05	-1.0E-05	2.3E-05	-4.4E-05	1.8E-04	1.1E-05	-5.3E-05
6X	-8.0E-06	-1.5E-05	-1.8E-05	-1.4E-05	-1.6E-05	2.2E-05	4.0E-06	2.8E-05	-1.5E-05	1.1E-05	6.8E-05	-2.1E-06
6Y	-2.4E-05	-1.6E-05	1.8E-05	2.8E-05	-1.5E-05	-6.2E-05	9.0E-06	8.0E-06	8.9E-06	-5.3E-05	-2.1E-06	1.3E-04
7X	1.4E-05	-2.4E-05	1.3E-06	-1.6E-05	1.5E-05	-1.3E-05	1.1E-05	7.9E-06	-1.8E-05	2.8E-05	-9.2E-06	-1.3E-05
7Y	7.6E-05	2.8E-05	5.8E-06	2.8E-05	3.9E-05	-5.6E-05	1.2E-05	-3.3E-05	-1.4E-05	-3.2E-05	-4.9E-05	-2.4E-05
8X	2.7E-06	3.1E-05	1.8E-05	7.2E-06	-1.1E-05	-3.0E-06	-3.3E-05	-1.1E-05	-1.5E-05	-1.7E-05	-2.6E-05	8.1E-06
8Y	-5.4E-05	7.8E-06	-9.1E-06	4.1E-06	1.1E-05	1.0E-05	-1.0E-06	-2.0E-05	2.3E-05	-2.4E-05	6.6E-06	-2.2E-05
9X	-2.6E-05	2.8E-05	-4.9E-06	7.3E-06	-4.5E-05	-1.6E-05	-1.5E-05	3.9E-06	-4.0E-07	2.6E-05	2.0E-05	1.1E-05
9Y	-6.1E-05	-1.6E-05	5.9E-06	2.3E-05	1.1E-06	-2.7E-05	1.3E-05	-1.4E-05	2.8E-05	-4.3E-05	1.2E-05	9.9E-06

	7X	7Y	8X	8Y	9X	9Y
1X	1.4E-05	7.6E-05	2.7E-06	-5.4E-05	-2.6E-05	-6.1E-05
1Y	-2.4E-05	2.8E-05	3.1E-05	7.8E-06	2.8E-05	-1.6E-05
2X	1.3E-06	5.8E-06	1.8E-05	-9.1E-06	-4.9E-06	5.9E-06
2Y	-1.6E-05	2.8E-05	7.2E-06	4.1E-06	7.3E-06	2.3E-05
3X	1.5E-05	3.9E-05	-1.1E-05	1.1E-05	-4.5E-05	1.1E-06
3Y	-1.3E-05	-5.6E-05	-3.0E-06	1.0E-05	-1.6E-05	-2.7E-05
4X	1.1E-05	1.2E-05	-3.3E-05	-1.0E-06	-1.5E-05	1.3E-05
4Y	7.9E-06	-3.3E-05	-1.1E-05	-2.0E-05	3.9E-06	-1.4E-05
5X	-1.8E-05	-1.4E-05	-1.5E-05	2.3E-05	-4.0E-07	2.8E-05
5Y	2.8E-05	-3.2E-05	-1.7E-05	-2.4E-05	2.6E-05	-4.3E-05
6X	-9.2E-06	-4.9E-05	-2.6E-05	6.6E-06	2.0E-05	1.2E-05
6Y	-1.3E-05	-2.4E-05	8.1E-06	-2.2E-05	1.1E-05	9.9E-06
7X	6.9E-05	4.0E-05	-3.8E-05	-1.7E-05	-4.5E-05	7.7E-06
7Y	4.0E-05	2.6E-04	-4.3E-05	-1.1E-04	-6.6E-05	-6.0E-05
8X	-3.8E-05	-4.3E-05	1.1E-04	3.8E-05	-7.2E-06	-1.0E-05
8Y	-1.7E-05	-1.1E-04	3.8E-05	1.1E-04	3.2E-06	4.1E-05
9X	-4.5E-05	-6.6E-05	-7.2E-06	3.2E-06	1.2E-04	3.7E-06
9Y	7.7E-06	-6.0E-05	-1.0E-05	4.1E-05	3.7E-06	8.6E-05

had coefficients drawn from a normal distribution with a standard deviation of 0.426, but whose mean depended on the value of the phenotype. When the phenotype fell in the negative direction from the starting value on a particular axis, then the mean of the $\beta\mathbf{H}$ was positive, thus tending to push the phenotype back towards its original form. Drift was simulated using a $\beta\mathbf{H}$ distribution centred on zero with a standard deviation equal to each eigenvalue divided by N_e. Each of these distributions is explained in more detail with each simulation. Note that a different random coefficient was drawn for each dimension at each generation in all simulations. Also note that the same mode was applied to all eigenvectors in each simulation; in principle, one could apply drift to some dimensions, directional selection to others, and stabilizing selection to some in any combination. None of the simulations can be considered deterministic or to be based on a "constant" rate or direction of evolution

Table 2. The eigensystem for the phenotypic space defined by P from the Bialowieza sample of *Sorex araneus*. The nine two-dimensional superimposed landmarks correspond to a 15-dimensional space. Eigenvalues and the percent of the total variance accounted for by each axis are in the first two rows. In the simulations, change on each axis was weighted by the corresponding eigenvalue. The loadings of the landmarks are shown in the next rows. Landmark numbers correspond to Figure 3.

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9
Eigenvalues	4.3E-04	3.7E-04	2.0E-04	1.9E-04	1.5E-04	1.3E-04	8.3E-05	7.0E-05	5.2E-05
% Explained	23.6%	20.5%	11.0%	10.6%	8.2%	7.4%	4.6%	3.9%	2.9%
1X	-0.29	-0.22	0.09	-0.20	-0.27	0.14	0.00	-0.22	0.20
1Y	-0.04	0.05	0.13	-0.40	0.31	0.12	0.19	-0.27	0.12
2X	-0.01	0.11	0.07	0.01	-0.12	-0.27	-0.32	0.48	-0.04
2Y	-0.02	0.38	0.06	-0.02	-0.06	0.09	-0.43	0.17	-0.07
3X	-0.11	0.01	-0.33	0.04	0.11	-0.17	0.21	0.02	-0.47
3Y	0.10	-0.47	-0.30	-0.15	-0.37	0.35	0.03	0.36	-0.13
4X	-0.03	0.05	-0.14	0.21	0.03	0.12	0.17	-0.01	-0.16
4Y	0.05	-0.16	0.16	0.21	-0.31	-0.02	-0.15	-0.16	0.33
5X	0.09	0.20	-0.19	0.02	0.14	0.30	0.39	0.30	0.49
5Y	0.01	-0.55	0.21	0.16	0.39	-0.38	0.10	0.19	-0.01
6X	0.14	-0.12	0.07	0.23	-0.10	0.25	-0.19	-0.47	-0.31
6Y	0.07	0.31	0.35	0.23	-0.40	-0.19	0.56	-0.01	-0.16
7X	-0.17	-0.07	-0.17	0.32	0.07	-0.27	-0.11	-0.18	0.36
7Y	-0.75	0.15	-0.01	-0.07	0.14	0.11	-0.07	0.01	-0.13
8X	0.14	0.06	0.06	-0.59	-0.19	-0.42	-0.04	-0.07	0.05
8Y	0.37	0.07	-0.36	-0.20	0.16	-0.10	-0.04	-0.23	-0.09
9X	0.24	-0.02	0.54	-0.03	0.34	0.33	-0.11	0.15	-0.12
9Y	0.23	0.22	-0.23	0.24	0.13	0.00	-0.19	-0.06	0.15

	PC 10	PC 11	PC 12	PC 13	PC 14	PC 15
Eigenvalues	4.3E-05	3.6E-05	2.8E-05	2.0E-05	6.0E-06	1.7E-07
% Explained	2.4%	2.0%	1.6%	1.1%	0.3%	0.0%
1X	-0.21	0.28	0.02	-0.20	0.21	0.55
1Y	-0.09	-0.24	-0.10	0.21	-0.14	0.14
2X	0.01	-0.25	-0.13	0.44	-0.10	0.39
2Y	0.11	0.45	0.28	-0.17	0.18	-0.07
3X	0.27	-0.09	-0.33	-0.39	0.22	0.08
3Y	-0.23	-0.07	-0.07	0.06	0.02	-0.24
4X	-0.13	0.25	0.23	-0.07	-0.78	0.05
4Y	0.55	-0.09	-0.32	-0.20	-0.30	-0.02
5X	0.33	-0.05	0.21	0.07	0.20	-0.04
5Y	0.08	0.10	0.37	-0.04	0.09	0.09
6X	0.18	-0.26	0.36	0.30	0.20	-0.07
6Y	-0.19	0.05	-0.04	0.12	0.15	-0.02
7X	-0.33	0.18	-0.24	0.20	0.13	-0.41
7Y	0.07	-0.15	0.00	0.10	-0.04	-0.20
8X	0.00	-0.15	0.24	-0.21	-0.10	-0.37
8Y	0.11	0.39	-0.21	0.32	0.01	0.14
9X	-0.14	0.09	-0.37	-0.15	0.02	-0.18
9Y	-0.41	-0.44	0.09	-0.40	0.03	0.19

because the coefficients were randomly selected from a range of values that changed direction and magnitude each generation.

Procrustes distance was used as a measure of divergence of shapes from the starting configuration. Procrustes distance is the square root of the sum of squared differences in the positions of the landmarks in two shapes (Slice *et al.* 1996; Dryden and Mardia 1998). Procrustes distance is a Mahalonobis distance that represents the overall

difference in the phenotype. The same Procrustes distance can describe the difference between many landmark configurations (Rohlf and Slice 1990). Though disadvantageous in some contexts, this property is useful because the distribution of distances reveals some aspects of the mode of evolution.

Changes in the phenotype are illustrated in some animations as thin-plate spline (*tps*) deformation grids showing the difference between the evolved shape and the starting configuration. The grids were constructed using techniques described in Bookstein (1989).

RESULTS

Simulation 1: Randomly Fluctuating Selection

Randomly fluctuating selection, in which the direction and intensity of selection changes at each generation, was modelled in the first simulation. This mode of evolution occurs when morphological fitness is influenced by many independent factors (e.g., mate choice, nutrition, winter temperature, predator density), each of which vacillates from year to year with changing environments and functional contexts. Randomly fluctuating selection is a type of random walk, or Brownian motion, because the direction and magnitude of change in any given generation is not influenced by that in earlier or later ones.

The randomly fluctuating mode is equivalent to evolution with a flat adaptive landscape, but one that has stochastic ripples in its surface (Figure 4). The landscape is flat because there is no particular direction in which fitness increases, but its surface ripples because selection is constantly pushing the morphology in one direction or another. This mode is *not* an example of genetic drift or neutral evolution. In drift, the change between generations is due solely to chance and is purely a function of variance and population size, whereas the selection coefficients applied here were larger than expected from drift. A simulation of genetic drift would have an adaptive surface that is flat and still because selection would not affect the morphology at all (see below). See Arnold et al. (2001) for a recent introduction to adaptive landscapes and multivariate evolution.

Selection coefficients for each of the 15 dimensions of the shape space were drawn each generation from a normal distribution with a mean of zero and a standard deviation of 0.426. These coefficients have an equal probability of being positive or negative and a mean absolute magnitude of 0.338. The simulation was run 100 times for 1,000,000 generations each run.

Figure 4. Adaptive landscape associated with randomly fluctuating selection. The overall surface is neither curved nor tilted, but random changes in fitness cause the surface to ripple. This selection moves the morphology in random directions at each step of the simulation. See an animation of the stochastic changes in the surface.

In the randomly varying mode of evolution, the 100 evolving shapes diffused randomly through the simulation space in an ever-enlarging cloud (Figure 5). Even though the mean rate of evolution was lower than that measured in *Cantius*, it was high enough so that in many of the runs the tooth shape evolved into functionally nonviable forms, as illustrated by the particular run shown in the left panel. A narrower distribution of selection coefficients (i.e., a smaller mean rate of per-generational change) would not have produced such extreme shapes. The evolution of nonsense molar shapes suggests that the **P** matrix is incompatible with a rate of evolution this high under this mode of selection, implying (1) that the true rate of molar evolution in shrews is lower; (2) that true rate of evolution is of a similar magnitude, but the **P** matrix changes to accommodate this rate and still maintains a functional tooth; or (3) that stabilizing selection removes the nonviable morphologies in the course of real evolution.

The most important result of this simulation is the distribution of morphological distances relative to the number of generations elapsed (Figure 5C). Procrustes distance from the ancestral form increased curvilinearly as a function of the square root of the number of generations elapsed. As the simulation proceeded, larger divergences in shape became more probable, while the likelihood of no divergence decreased. Put another way, we may expect multivariate morphological shape to diverge from its ancestral condition in a curvilinear fashion under random fluctuating selection. This pattern

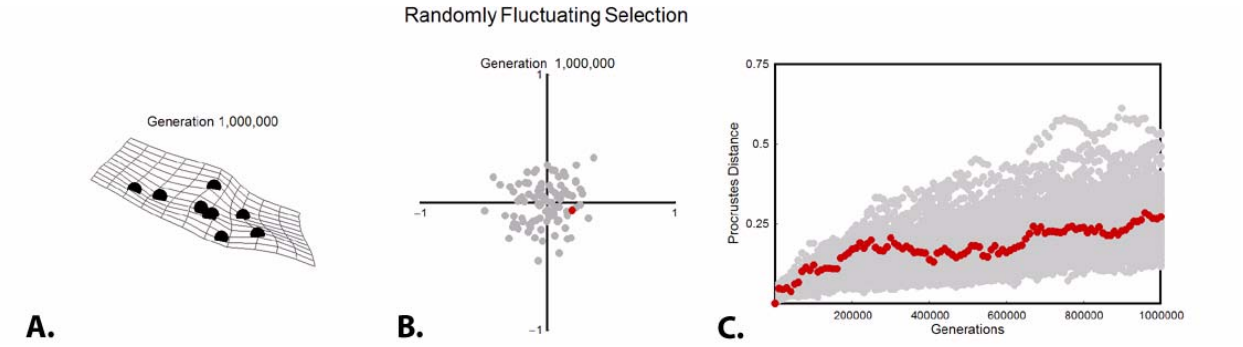

Figure 5. The results of randomly fluctuating selection. **A.** The end shape of one of the 100 runs as landmarks and a deformation grid. **B.** The positions of the 100 end shapes on the first two axes of the principal components space. The red dot is the position of the shape shown at left. Only the first two components are illustrated, but the results in panels A and C are based on all 15 axes. **C.** Divergence from the starting morphology of all 100 runs in Procrustes units. The red trace shows divergence in the morphology pictured in the left panel. **D.** Animation of the entire simulation sampled every 10,000 generations. The period of time covered by this simulation is 1,000,000 generations, which in shrews is equivalent to 800,000 – 1,000,000 years, approximately half the duration of the Pleistocene.

has implications for reconstructing phylogeny from morphometric shape and for estimating divergence times based on morphological differences, as discussed below.

Simulation 2: Directional Selection

Directional selection, in which selection more often causes changes in one direction than another, was modelled in the second simulation. The preponderance of selection coefficients was in the negative direction, but with a stochastic component to both direction and magnitude at each generation. This mode of evolution is equivalent to an adaptive landscape that slopes ever upwards so that fitness is generally higher in the negative direction, but with stochastic fluctuation creates short-lived, small scale increases in fitness in other directions (Figure 6). The tilt of this landscape is determined by the degree of positive or negative bias in the selection coefficients, which was chosen arbitrarily for this simulation.

To carry out this simulation, selection coefficients were drawn from a Gamma distribution whose α and λ parameters were 100 and 0.1

Figure 6. The adaptive landscape associated with directional selection. The overall tilt pushes evolution in the negative direction along each PC axis, but stochastic fluctuations in the surface may force the morphology in any given direction in the short term. You can see an animation of the stochastic fluctuations in the surface.

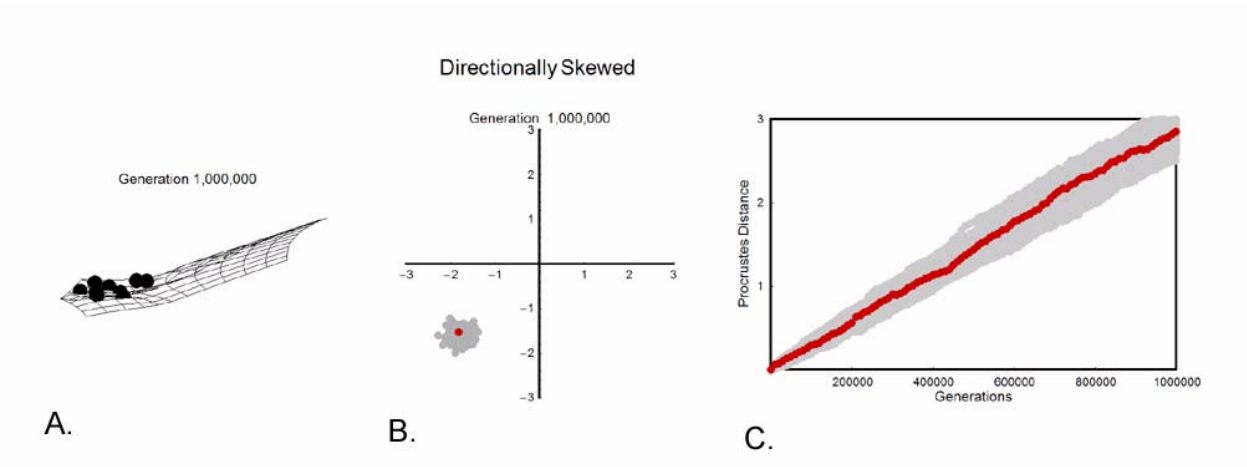

Figure 7. Results of directional selection. **A.** The final phenotype of one of the 100 runs. **B**. The final distribution of all 100 runs in the first two PCs. Note that all have moved in a negative direction along both axes. The red dot indicates the position of the phenotype shown in panel A. Only the first two components are illustrated, but the results in panels A and C are based on all 15 axes. **C.** Distribution of Procrustes distances in relation to time. The red trace indicates the distances of the phenotype shown in A. **D.** Animation of the full simulation in 10,000 generation intervals. The period of time covered by this simulation is 1,000,000 generations, which in shrews is equivalent to 800,000 – 1,000,000 years, approximately half the duration of the Pleistocene.

respectively, scaled to have a mean of zero and a standard deviation of 0.426. A Gamma distribution is skewed in one direction or the other, and coefficients drawn from this one have the same mean intensity as in Simulation 1, but 54% of them are negative. The simulation was run 100 times for 1,000,000 generations each run.

The results show a rapid rate of morphological change that is clearly directional whether viewed as a morphological reconstruction, as the cloud of points in phenotypic space, or the change in Procrustes distance with time (Figure 7). Even though the average magnitude of selection at any given generation was the same as in the first simulation, the magnitude of morphological change over the entire simulation was more than six times higher as measured in Procrustes distances. The consistent changes in the same direction throughout the simulation cause this apparently rapid and dramatic change, whereas in the randomly fluctuating mode the morphology often doubled back in a homoplastic manner, making the greatest changes from the ancestral form much smaller. The rate of directional change depends on the proportion of negative to positive coefficients. Had the proportion of negative values been less than 54%, then the departure from the ancestral form would have been slower.

Simulation 3: Stabilizing Selection

Stabilizing selection, in which selection pushes the morphology back towards the ancestral form if it moves too far away, was modelled in the third simulation. Stabilizing selection keeps a population at an optimum (Schmalhausen 1949). The intensity of selection increases as the population

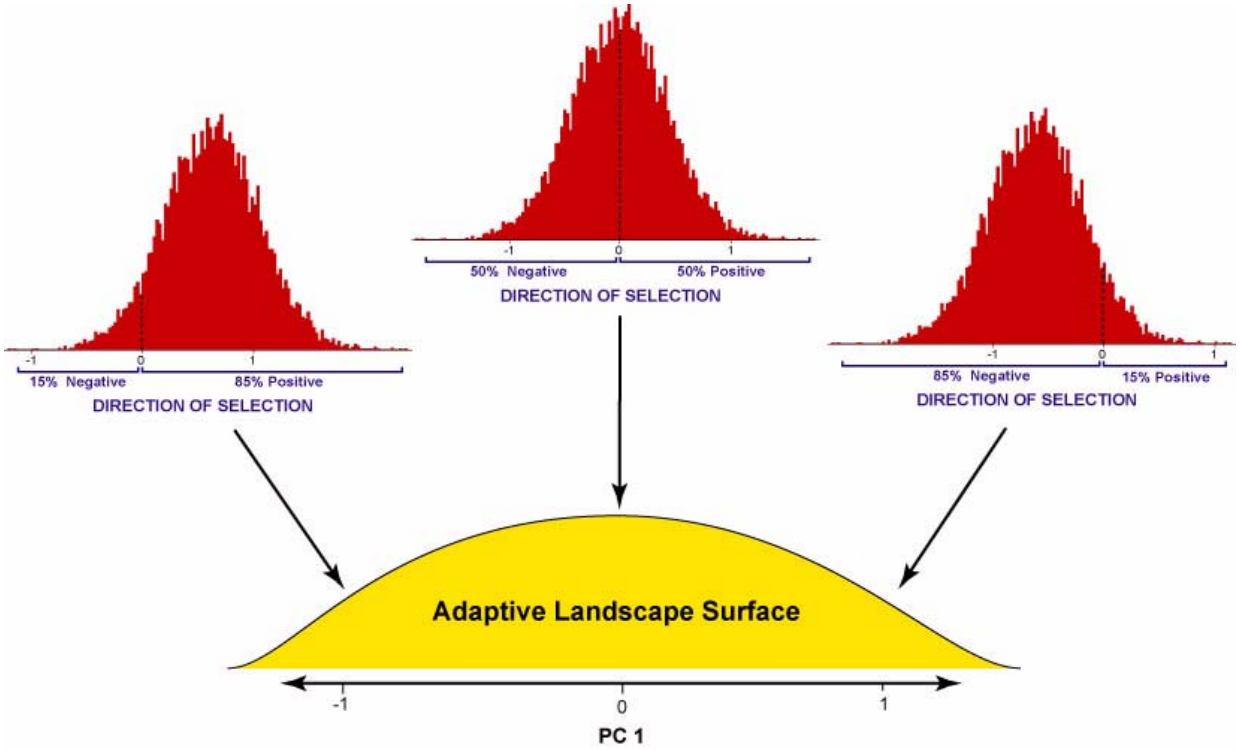

Figure 8. The adaptive landscape surface and distributions of selection coefficients used to simulate stabilizing selection. The height of the yellow surface indicates the relative fitness of the phenotype in variation space. The starting shape, which is located at the centre of the peak, is the most fit. At the peak, the distribution of coefficients has an equal proportion of positive and negative values. As the phenotype moves off the peak to the negative side, the distribution of coefficients has a greater proportion that are positive; when the phenotype moves in a positive direction, the coefficients become proportionally more negative.

mean moves away from the optimum, culling morphologies that are radically different. In palaeontology, stabilizing selection is one of several processes thought to explain patterns of morphological stasis (Vrba and Eldredge 1984; Maynard Smith et al. 1985; Lieberman et al. 1995).

Stabilizing selection was simulated here using a single-peaked adaptive landscape to adjust the distribution from which selection coefficients were drawn (Figure 8). The peak represents the optimal form, and the slopes represent morphologies that are less fit. Mathematically, the adaptive landscape surface and corresponding selection distributions are described by Lande (1976) and Arnold et al. (2001).

The adaptive surface is a Gaussian normal fitness curve whose width is 0.85 Procrustes units. When the simulated morphology is at the peak, selection coefficients are drawn from a normal distribution with a mean of zero and a standard deviation of 0.426, precisely the same distribution as in the simulation of randomly fluctuating selection. The width of the peak is about three times larger than the average magnitude of the selection coefficients when the morphology is at the peak (0.29 Procrustes units). Because of this small proportional difference, the stabilizing selection in this simulation is quite strong.

As the simulated morphology moves off the peak, the mean of the distribution of coefficients is shifted in an opposite direction so that the chance of a positive coefficient increased as the morphology moves in a negative direction and vice versa. Mathematically, the following fitness function was used

$$W(z) = \frac{Sign[z] \cdot z^2}{2 \cdot .85} \quad ,(3)$$

in which *W(z)* is the fitness of the phenotype at position *z* on the peak. Position of zero is the top of the peak. This equation was translated into a modification of the selection coefficients by adding *W(z)* to the mean of the distribution from which they were drawn. Thus, when the morphology moves in a negative direction, the mean of the distribution increases and pushes the morphology back in a positive direction. The random nature of the distribution of coefficients adds a stochastic

Figure 9. The adaptive landscape represented by the simulation of stabilizing selection. The optimum is at the centre of the peak, and fitness decreases away from that peak. Note the small fluctuations in the surface, which are due to the stochastic nature of the selection coefficients. The 100 phenotypes are shown as black and red dots on the surface. The red dot indicates the phenotype illustrated in Figure 10A. An animation of the fluctuating surface can be viewed.

component to the direction and magnitude of selection (Figure 9).

The results of this simulation showed a rapid morphological divergence from the mean during the first few thousand generations but one that quickly stabilized at a fixed mean from the ancestral form (Figure 10). Stabilizing selection kept the morphology stable, and the diffusion through principal components space was small and concentrated. Most notably, the Procrustes distances of the 100 runs remained very small.

Simulation 4: Genetic Drift (Neutral Evolution)

Genetic drift, or random change in the population mean due to generational sampling in the absence of selection, was modelled in the fourth simulation. Like random selection, genetic drift is a type of random walk. The difference between drift and random selection is that the magnitude of change at each generation is much smaller in the former than the latter. In drift, the rate of per-generation change is a function of the phenotypic variance (literally the heritable variance) and population size (Wright 1931; Lande 1976). Drift is equivalent to evolution on a flat, placid adaptive landscape (Figure 11). The phenotype is neither more nor less fit, no matter which way it moves, and there is no stochastic selection to create transient ripples in the surface.

The amount of change in each generation depends only on the phenotypic variance and long-term effective population size (N_e), with change being smaller when the population is larger. I ran two simulations of drift because of the discrepancy between molecular and field estimates of population size in *Sorex araneus*: one used N_e = 70,000 and the other N_e = 70.

When N_e was 70,000, absolutely no change was visible in the phenotype at any time during the simulation (Figure 12). The 100 runs diffused through the principal components space with the same pattern as in randomly fluctuating selection, but the average distance travelled was about five orders of magnitude smaller. When N_e was dropped to 70, some phenotypic change was visible (Figure 13), but considerably less than under stabilizing selection. This result suggests that the amount of change in molar morphology due to drift is very, very small, even over 1,000,000 generations (roughly equivalent to 1,000,000 years).

Simulation 5: Randomly Fluctuating Selection with Speciation

Randomly fluctuating selection with speciation was modelled in the final simulation. The mode of evolution in this simulation was the same as in the first, but instead of treating each run as a single evolving lineage, the lineage branched twice, once at the beginning of the simulation and then again

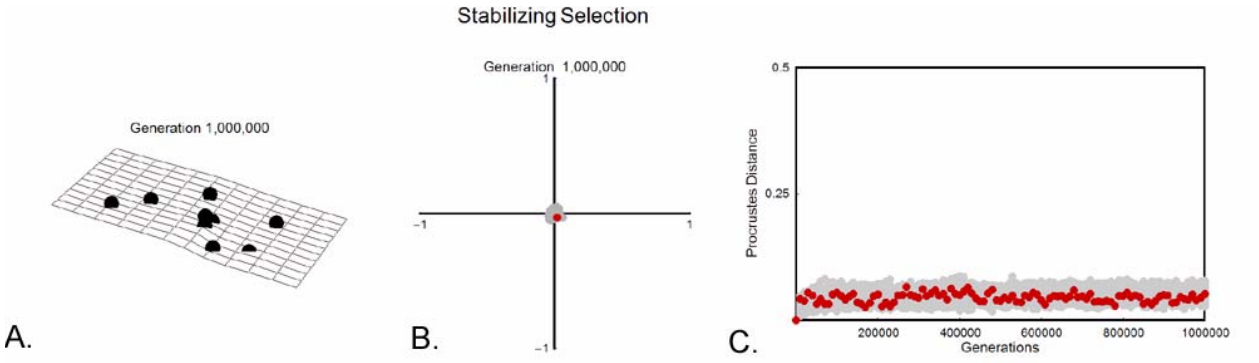

Figure 10. Results of stabilizing selection. **A.** The final phenotype of one of the 100 runs after 1,000,000 generations. Note that the phenotype is very similar to the starting phenotype shown in Figure 3. **B.** The distribution of all 100 phenotypes in principal components space after 1,000,000 generations. Note the very small dispersion compared to the first simulation. The red dot indicates the phenotype illustrated in A. Only the first two components are illustrated, but the results in panels A and C are based on all 15 axes. This panel is a two-dimensional version of Figure 9. **C.** The distribution of Procrustes distances relative to generation. The red trace indicates the distance of the run culminating in A. **D.** Animation of the entire simulation. The period of time covered by this simulation is 1,000,000 generations, which in shrews is equivalent to 800,000 – 1,000,000 years, approximately half the duration of the Pleistocene.

halfway through (Figure 14). Each simulation was run 100 times for 1,000,000 generations.

This simulation shows how derived morphological similarity accumulates within clades (Figure 15). During the first 500,000 generations, the two stem lineages diverge randomly from one another, creating two derived ancestral morphologies for their respective daughter species. By the end of the simulation, those shared derived phenotypes are still visible, even though each terminal phenotype has its own autapomorphic condition. Shared phylogenetic history also leaves its trace in the positions of the phenotypes within the PC space. The daughter species of the first clade (red dots) remained relatively close to one another but distant the daughters of the second (blue dots).

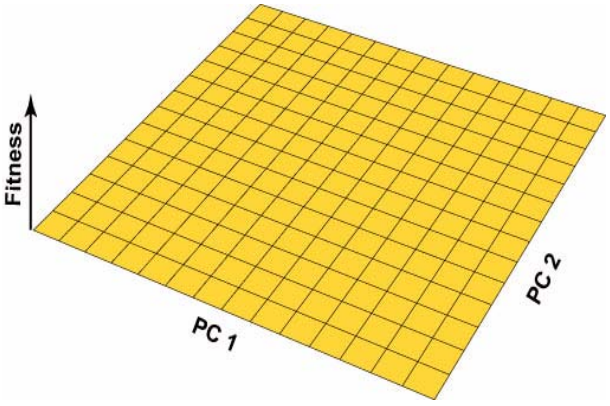

Figure 11. The adaptive landscape associated with genetic drift. The surface is completely flat with no fluctuations. Movement of the phenotype is due to sampling error from one generation to the next, and depends only on the phenotypic variance and effective population size.

DISCUSSION

The Comparative Imprint of Different Modes of Evolution

Directional, stabilizing, and randomly fluctuating selection each leave a distinctive imprint on the distribution of morphological distances, a phenomenon that is well known for univariate traits (e.g., Gingerich 1993; Hansen and Martins 1996; Roopnarine 2001; Figure 16). The three modes influence the scaling relationship between short- and

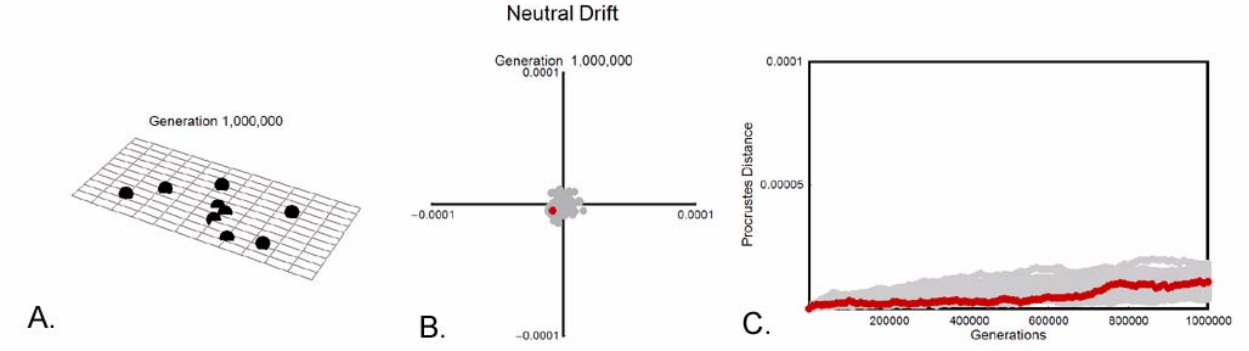

Figure 12. The results of genetic drift when N_e = 70,000. **A.** The final phenotype of one of the 100 runs after 1,000,000 generations. No difference is visually detectable between this and the starting shape. **B.** Distribution of all 100 phenotypes on the first two axes of principal components space after 1,000,000 generations. Note that the scale of the axes is four orders of magnitude smaller than in Simulation 1. The red dot indicates the position of the phenotype illustrated in A. Only the first two components are illustrated, but the results in panels A and C are based on all 15 axes. **C.** The distribution of Procrustes distances in respect to generation. **D.** Animation of the entire simulation. The period of time covered is 1,000,000 generations, which in shrews is equivalent to 800,000 – 1,000,000 years, or approximately half the duration of the Pleistocene.

long-term evolutionary divergences in different ways, and the patterns can be used to infer the average evolutionary mode responsible for the differences among phenotypes.

In directional selection, the rate of divergence over very short intervals is equal to very long intervals, and the morphological distance from the ancestor increases linearly with time. Morphological distance is a straight, diagonal line when plotted against time since divergence.

In randomly fluctuating selection, divergence from the ancestral form is greater over short intervals than over long ones, so that morphological distance increases curvilinearly with time. Drift leaves exactly the same imprint and randomly fluctuating selection, but the rate of divergence in the former is much slower. It is not correct that evolution in these modes "slows down" (Kinnison and Hendry 2001; Sheets and Mitchell 2001) – after all, we know that the rate of change at each generation remains the same on average throughout the simulation because the distribution of selection coefficients in the simulation is kept constant – but, rather, the probability that selection will take the shape at least some distance back towards its ancestral form increases with the amount of divergence (Berg 1993; Gingerich 1993). In other words, the chance increases with time that a long run of selection coefficients of the same sign will take the shape towards the ancestral form rather than away from it. The overall effect is that the net evolutionary divergence is greater over short intervals than long intervals, and the expectation of morphological distance is a function of the square root of the number of generations since common ancestry. Morphological distance forms a curved line when plotted against time since divergence.

In stabilizing selection, the first few thousand generations look like the randomly fluctuating pattern, but divergence then settles down to a roughly

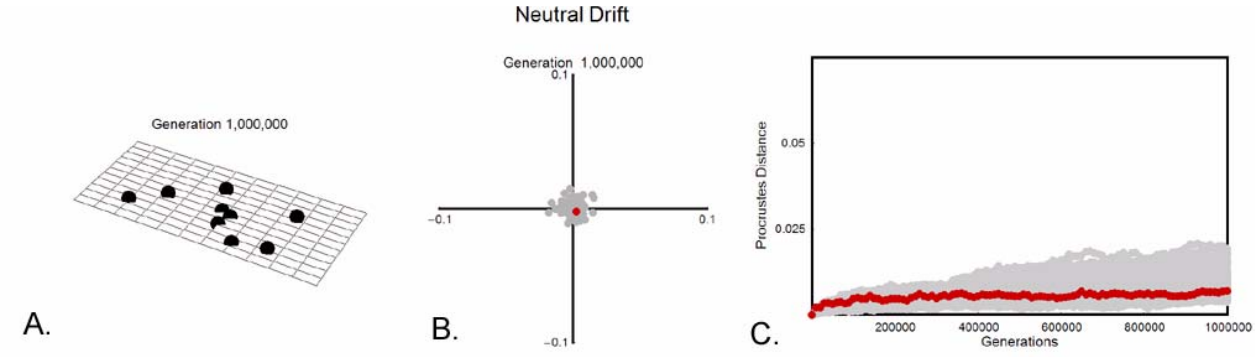

Figure 13. The results of genetic drift when N_e = 70. **A.** The final phenotype of one of the 100 runs after 1,000,000 generations. Only a slight difference is visually detectable between this and the starting shape. **B.** Distribution of all 100 phenotypes on the first two principal components after 1,000,000 generations. Note that the scale of the axes is one order of magnitude smaller than in Simulation 1. The red dot indicates the position of the phenotype illustrated in A. Only the first two components are illustrated, but the results in panels A and C are based on all 15 axes. **C.** The distribution of Procrustes distances in respect to generation. **D.** Animation of the entire simulation. The period of time covered by this simulation is 1,000,000 generations, which in shrews is equivalent to 800,000 – 1,000,000 years, approximately half the duration of the Pleistocene.

constant distance. Morphological distance is distributed as a straight, horizontal line when plotted against time since divergence. Note that with the parameters used here, the rate of evolution over short intervals is faster than over a similar interval under neutral drift, but total divergence is less over long intervals. Many tests of evolutionary mode are based on the assumption that neutral evolution is intermediate in rate between directional and stabilizing selection, regardless of the interval measured (Turelli et al. 1988; Lynch 1990; Spicer 1993). The relation between the two modes will depend on the curvature of the stabilizing adaptive surface, as well as on the existence and magnitude of transient stochastic selection on the surface.

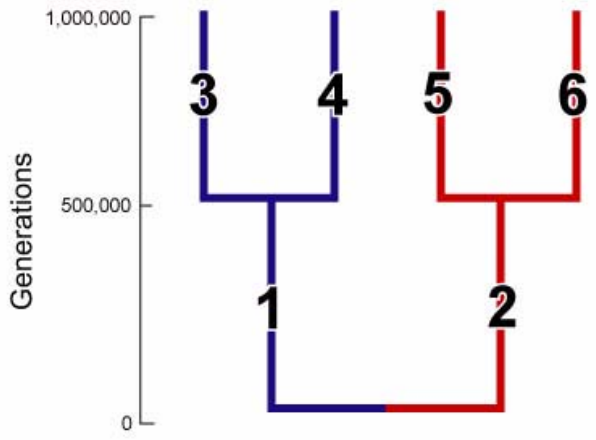

Figure 14. The structure of the clades used to simulate phylogenetic divergence in shape. For the first 500,000 generations, two daughter species diverge from a common ancestor; after 500,000 generations each of those split into two daughters.

What Imprint Do Real Data Have?

Morphological distances among real shrew molars are distributed as though they evolved under randomly fluctuating selection (Figure 17). In this graph, morphological distance in molar shape is plotted against mtDNA sequence divergence, which serves as a proxy – albeit an imperfect one – for time since common ancestry. Each point is a pairwise comparison between two taxa: some populations within the same species, some congeneric species, and others species belonging to different genera (see Polly 2003a for more

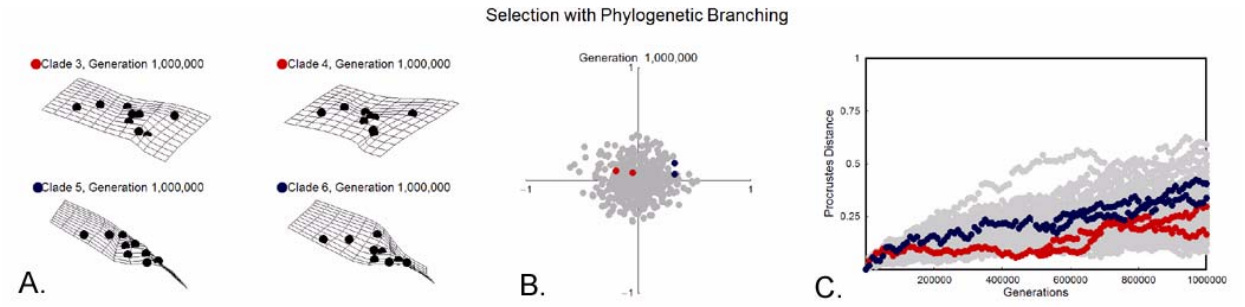

Figure 15. The results of randomly fluctuating selection with a branching phylogeny. **A.** The phenotypes of the four terminal populations of one of the runs. Note the derived similarities shared by the two members of each clade. **B.** Position of all 400 phenotypes (100 runs times 4 terminal clades) in the first two PCs of the simulation space. Clade 3 plus 4 are shown by red dots and 5 plus 6 by blue. Note that the positions of the two members of each clade are relatively close to one another. Only the first two components are illustrated, but the results in panels A and C are based on all 15 axes. **C.** Divergence of all phenotypes from the starting shape. The red trace shows clade 1 plus 3 plus 4 and the blue trace clade 2 plus 5 plus 6. Note this graph does not show the Procrustes distances among the clades, only between each descendant phenotype and the ancestral one. **D.** Animation of the entire simulation. The period covered by the simulation is 1,000,000 generations, which in shrews is equivalent to 800,000 – 1,000,000 years, or approximately half the duration of the Pleistocene.

details). Morphological distance between the two is measured as Procrustes distance between the same nine landmarks used in the simulations. Phylogenetic divergence is measured as a Kimura 2-parameter distance in the cytochrome *b* mtDNA sequences of the two taxa. This genetic distance is expressed as the percentage by which the two sequences differ, taking into account the probability of homoplasy at any position within the sequence. Generally, this measure of genetic difference increases linearly with time since common ancestry – though chance and homoplasy obscure this pattern – and so can be used as a coarse measure of time since common ancestry when direct palaeontological estimates are not available (Brown et al. 1979; Springer 1997). A second axis is shown with the genetic distances converted into millions of years using a rate of divergence of 2% per million years (Brown et al. 1979; Klicka and Zink 1997, 1998). Shrews are not well-studied palaeontologically, but coarse estimates of divergence times among the same taxa can be made from the fossil record (Repenning 1967; Harris 1998; Rzebik-Kowalska 1998; Storch et al. 1998). This comparison suggests that molecular clock estimates may be overestimating divergence times, but in any event the last common ancestor of the most distantly related *Sorex* taxa was at least 15 million years ago sometime during the Middle Miocene. The x-axis of the figure is thus 15 or more times longer than the simulations in Figure 16.

The real data have the same curvilinear pattern of morphological divergence found in randomly fluctuating selection and drift. The pattern is neither linear, like directional selection, nor flat like stabilizing selection. Furthermore, the rate of divergence is too high for drift, ruling out that mode as well. In the real data, divergences of one million

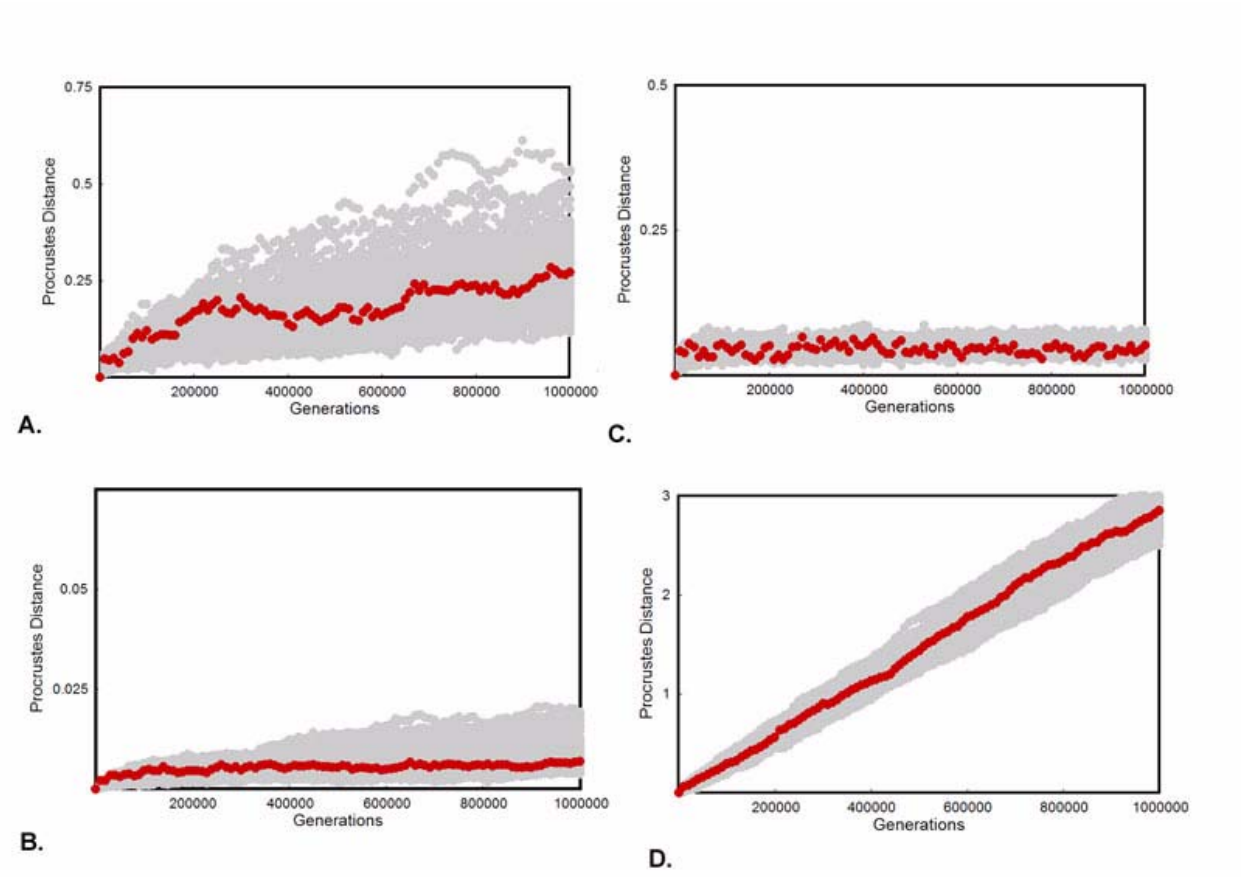

Figure 16. Comparison of morphological divergence under different evolutionary models. **A.** Randomly fluctuating selection. **B.** Random drift (N_e = 70). **C.** Stabilizing selection. **D.** Directionally biased selection.

years are roughly the same as 2-3% sequence divergence and have associated Procrustes distances of 0.05 to 0.75. Even in the less conservative drift simulation that used N_e= 70, the largest Procrustes distances reached over 1,000,000 generations were less than 0.025. The real data have a much slower divergence than in the randomly fluctuating selection simulation, though, which reached morphological distances of 0.5 in only 1,000,000 generations, whereas the real data only have distances as high as 0.25 over 10 to 15 million years (generation lengths in shrews are roughly one year long). Clearly, the selections intensities driving molar shape evolution in shrews are lower than those measured in *Cantius*.

A statistical test for mode could be developed from the curvatures and slopes of best fit lines through the Procrustes distributions generated by the simulations. Such a test is beyond the scope of the present paper, but visual inspection is enough to indicate that the curvature of the line through the real data is enough to rule out either directional or stabilizing selection, and the amount of shape divergence is too great for drift.

While impossible to test here, it can be hypothesized that the effects of mutation and selection on the interlocking between occluding teeth explain both the higher-than-neutral but still relatively slow rate of change. Shrew molars are of the tribosphenic type, and have a tight, complicated fit between their three-dimensional cusps, blades, and basins (Butler 1961; Evans and Sanson 2003). This lock-and-key mechanism ensures that occluding teeth must evolve together (Polly et al. 2005). If a crown feature on one tooth changes, then functional selection will either remove that variant or favour a corresponding change on the counterpart tooth. Because of this functional constraint, tooth morphology cannot be free to drift (presuming that occluding teeth are under somewhat independent genetic control), but selection may not be directional and will be slow because any change will require a corresponding one in another tooth. Ultimately, these functional constraints will impose morphological boundaries beyond which evolution cannot pass. Several of these simulations pass those boundaries, most notably directional selection, creating nonsense

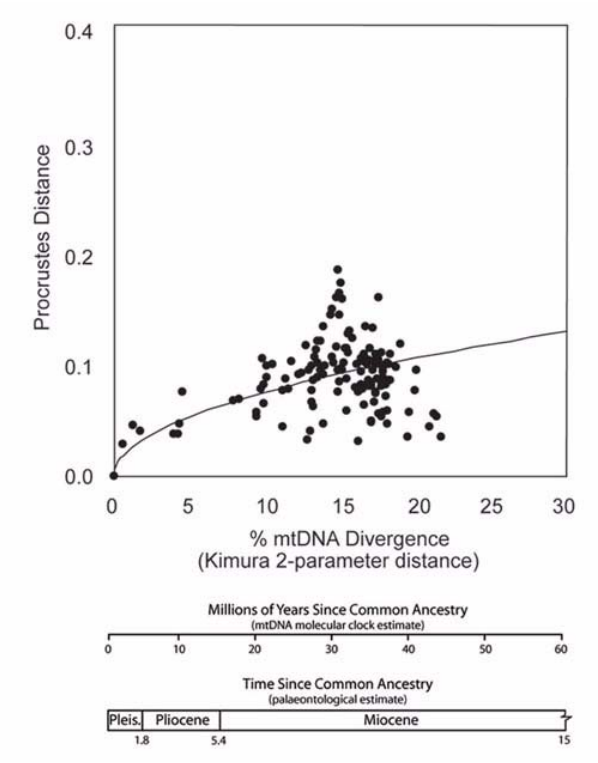

Figure 17. Morphological divergence in molar shape relative to phylogenetic divergence in extant shrews. Each point is a pairwise comparison between two taxa showing the morphological distance between their tooth shapes as Procrustes distance, and the phylogenetic distance between the two as measured by cytochrome *b* mtDNA. Two alternate timescales are provided for the x-axis: (1) millions of years since common ancestry as estimated using a mtDNA divergence rate of 2% per millions of years, and (2) period of common ancestry as estimated from palaeontological data. The pattern of morphological divergence corresponds to an evolutionary mode of randomly fluctuating selection (c.f., Figure 16). (Figure reproduced from Polly 2003a).

tooth shapes. If evolution were to approach those boundaries, stabilizing selection would prevent further change. The pattern of divergence among the real shrew taxa shows no sign that such boundaries are exerting an effect.

Phylogeny Reconstruction of Shape

An important feature common to all the evolutionary modes is that there is almost no chance that a derived morphology will be exactly like the ancestral one. In other words, the chance that homoplasy will recreate precisely the ancestral condition of a multivariate trait is nearly nil. This observation is important for phylogeny reconstruction, because it means that a predictable relationship exists between divergence in quantitative measures of multivariate morphology and time since common ancestry. In the directional selection, random selection and drift modes, a derived descendant morphology will be different from its ancestral condition and the amount of difference will have a linear relationship with the phylogenetic time elapsed. Multivariate quantitative phylogeny reconstruction algorithms are capable of recovering branching patterns when such conditions are obtained (Felsenstein 1988). Only in the case of strong stabilizing selection, in which morphology does not diverge from its ancestral form, is phylogeny not easily recoverable – even here, though, a problem with reconstructing phylogeny only arises when the boundaries of the stabilizing selection have been visited many times. For that period when lineages of a clade are wandering around the peak of an adaptive landscape without being driven back towards its centre, the pattern of divergence will be like that of fluctuating selection and phylogeny will be recoverable.

Univariate traits, such as trait size, do not have this expectation of divergence from the ancestral form. Regardless of the length of time, the most likely descendent value for a univariate trait is the same as its ancestor, even though the probability of a large difference increases with the number of generations. With univariate traits, the chance that homoplasy will recreate the precise ancestral condition is fairly high. The reason that multivariate traits, such as tooth shape, behave differently can be understood by thinking of them as a collection of several univariate traits: the probability that all of the traits will return to their ancestral values at the same time is small, even though each one has a high expectation of doing so; the more variables there are in the complex morphology, the less likely the chance that homoplasy will precisely recreate the ancestral form. (Incidentally, this observation explains the difference in rate scaling of univariate molar size (Polly 2001b) as compared to multivariate molar shape (Polly 2002) in viverravid carnivorans.)

The potential of morphometric shape for phylogeny reconstruction can be seen in Simulation 5 (Figure 15). The phenotypes within each clade share derived similarity not present in the ancestor or in the other. The red clade, for example, has a large, wide talonid basin at the posterior end of the tooth, while the blue clade has a short, narrow talonid. The shared derived similarity of each clade can be seen not only in the morphology itself, but also in the positions of the taxa within the principal components principal components space. In that space, any point away from the origin of the two axes represents a derived morphology, and the two

red dots lie in a common derived space separate from where the two blue dots lie. The position of the taxa within PC space is the key to reconstructing phylogeny from morphometric representations of shape. Each axis of PC space can be treated as an independent trait, and the score of each taxon on that axis as its trait value; the combination of values on the different axes will be unique for each taxon, but related taxa will share similar derived values different from zero on many of the axes. To reconstruct phylogenetic relationships, one must find the tree that best explains the positions of the taxa simultaneously across all of the PC axes (Polly 2003a, b). When the traits have evolved under a Brownian motion mode of evolution, which appears to be the case for shrew molar teeth, then a maximum-likelihood method for continuous traits would be an appropriate algorithm for phylogenetic analysis (Felsenstein 1973, 1981, 1988, 2002).

On "Morphological Clocks"

Complex morphology that evolves under randomly fluctuating selection, drift, or directional selection can be used as a coarse "morphological clock" in the same way that molecular sequence divergence is used (Polly 2001a). Because morphological distances increase in linear or curvilinear fashion under these modes of evolution, with them it is possible to predict the time since divergence of two taxa when one knows the Procrustes distance between their shapes and when one has an estimate of the per-generation rate of evolution. The principle is the same as with neutral molecular sequence divergence (Kimura 1983), though the relationship between morphological shape to time is not as tightly linear.

The error in morphological clocks, especially under randomly fluctuating selection, would be large, however. For example, the best estimate of time since divergence for two taxa with a molar Procrustes distance of 0.25 would be 600,000 years given the parameters used in simulating randomly fluctuating selection (Figure 16); however, divergences as small as 200,000 year can produce distances of 0.25, as can divergences well over 1,000,000 with the same parameters. The situation is better with directional selection, where a Procrustes distance of 1.0 implies a divergence time of about 375,000 years, with a range of possibilities from 220,000 to 400,000. In some cases, however, these estimates may be useful when no other data on divergence times are available. More often, though, direct stratigraphic estimates of common ancestry based on known fossil occurrences of taxa will probably be better than such "morphological clock" estimates.

P-matrix Effects and Limitations

The **P** matrix influenced the distribution of phenotypes that could evolve; some morphologies were likely, and some nearly impossible. These constraints can easily be seen in the phenotypic space (Figure 18). Each panel shows the 100 evolving phenotypes from the four main simulations. The effect of the correlations imposed by **P** can be most readily seen in the results of randomly fluctuating evolution. The positions of each landmark are shown, and a 95% confidence ellipse has been drawn around them. **P** prohibits, or at least renders improbable, those phenotypes that lie outside the trajectories of the expanding ellipses. Had there been no covariances among the landmarks, then the confidence ellipses would be perfectly circular and any phenotype would be possible given enough generations. The **P** matrix used in these simulations was estimated from a real population, and so the constraints on morphological evolution imposed by it are probably real. The correlations recorded in **P** come from many sources: developmental interactions, genetic pleiotropy, and chance sampling among them.

Limits exist to the amount of morphological evolution possible with a particular **P** matrix. Its correlations impose limits to how much a particular part of the morphology can change and still be biologically compatible with the others. This phenomenon is most evident in the results of directional selection (Figure 18B). The phenotypes move rapidly away from their ancestral form under directional selection, and after only 50,000 to 100,000 generations, landmarks have started to move into biologically incompatible positions. By 130,000 generations some points have fused, and by 170,000 begun to move past one another. From that point on, the simulated morphologies could not function as real teeth, demonstrating the limits that **P** imposes on diversification. In life, three possibilities may exist: (1) evolution progresses slowly without reaching the limits of **P**; (2) stabilizing selection prevents the phenotype from reaching those limits by pushing it back towards the ancestral form; or (3) **P** itself changes to accommodate new, functionally viable morphologies that are radically different from the ancestral form.

But does **P** remain constant in real biological systems? One of the assumptions of the simulations was that **P** did remain constant, but this issue is of major current concern in evolutionary genetics (Roff 1997; Lynch and Walsh 1998). Logically, phenotypic and genetic covariances must evolve given enough time because if they did not, all organisms would have morphologies of roughly the

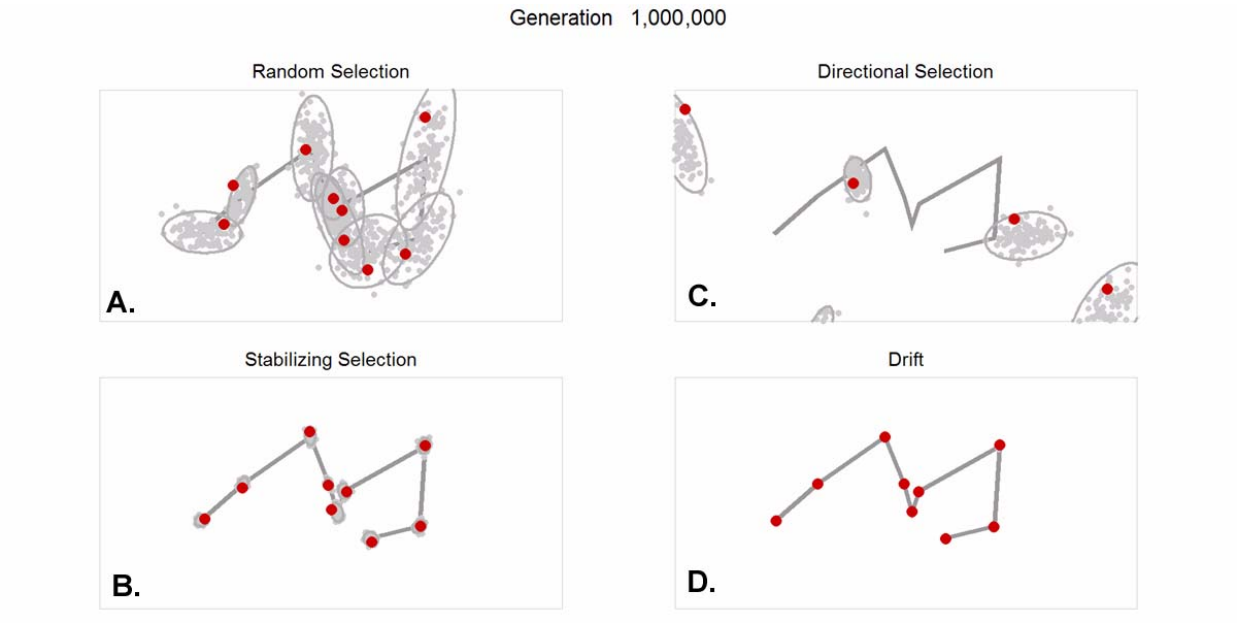

Figure 18. Simulations in phenotypic space. Each panel shows 100 phenotypes, each as nine landmarks, at the end of 1,000,000 simulation generations. Around each landmark is a 95% confidence ellipse, whose orientation helps show the pattern of covariances between landmarks. The ancestral phenotype is shown as a grey line. **A.** Phenotypes undergoing randomly fluctuating simulation. **B.** Phenotypes undergoing stabilizing selection. **C.** Phenotypes undergoing directional selection. **D.** Phenotypes undergoing drift (N_e = 70). **E.** Animation of the complete simulation.

same form (Lofsvold 1986). Genetic studies have demonstrated that showing that **P** (and **G**, the heritable phenotypic variances and covariances) can change significantly, even among closely related species (Lofsvold 1986; Kohn and Atchley 1988; Steppan 1997; Badyaev and Hill 2000; Roff et al. 2004). However, the changes in covariances, even when statistically significant, are often small, suggesting that trait correlations evolve slowly and remain broadly the same over long periods of phylogenetic divergence (Kohn and Atchley 1988; Brodie 1993; Steppan 1997; Arnold and Phillips 1999;

Roff et al. 2004). In shrew molars, **P** has been demonstrated to evolve over timescales as short as 10,000 years, but the effect of those changes on the overall covariance structure is small (Polly 2005). In principle, the simulation presented in this paper could be adapted to allow for a changing variance-covariance matrix.

CONCLUSION

Nearly 20 years ago Masatoshi Nei (1987) asserted that the study of evolution is focussed on two types of questions: reconstructing the histories of individual groups of organisms and understanding the mechanisms that drive evolutionary change. Nei's dichotomy, which he used to introduce his book on *Molecular Evolutionary Genetics*, served as a rhetorical device around which he integrated emerging knowledge of genetic diversity at the molecular level, concluding that the analysis of DNA and protein data would rapidly extend knowledge on both fronts. But more important for Nei was the possibility that phylogenetic and evolutionary inference might be fused through the quantitative analysis of molecular data, and his book was a bold attempt to achieve that synthesis. Now, two decades later, the power of Nei's molecular genetic synthesis can hardly be doubted. The expansion of molecular studies has far outstripped any other evolutionary genre, including palaeontology. Population-level data on geographic variation in gene sequences, for example, provide insights into mutation rates, demography, migration, speciation, phylogenetic diversification, times of common ancestry, and response to geologic-scale environmental restructuring (Patton and Smith 1989; Avise 1994, 2000; Wakeley and Hey 1998; Hewitt 1999, 2000; Knowles 2004). The simultaneous inference of both pattern and process from genetic data is now the norm.

Nei's dichotomy still occurs in palaeontology, however, where quantitative synthesis of evolutionary process and phylogenetics is still in its infancy. While the field is rich, both with studies of the histories of groups of organisms and studies of evolutionary mechanisms, the data and methods employed in the two types of study are radically different. Discrete-character, parsimony-based cladistics remains the dominant paradigm for phylogeny reconstruction in palaeontology (Adrain 2001; Padian et al. 1994), but quantitative analysis of morphological size and shape (Raup 1967; MacLeod and Rose 1993; Foote 1994, 1997, 1999; Jernvall et al. 1996; Smith 1998; Gingerich and Uhen 1998; Smith and Lieberman 1999; Roopnarine 2001) or of taxonomic abundance and extinction (Raup 1976; Raup and Marshall 1980; Sepkoski 1997; Alroy 2003; Janis et al. 2004) are the most common palaeontological approaches to the study of evolutionary mechanisms. Indeed, many palaeontologists share a widespread mistrust that quantitative measures of morphology do not carry recoverable information about phylogenetic and evolutionary history (see reviews in MacLeod 2002; Humphries 2002; Felsenstein 2002).

One barrier to the quantitative synthesis of evolutionary mechanisms and phylogenetics in palaeontology is the lack of tools for comparing the evolutionary behaviour of complex morphological traits (e.g., skulls, teeth, lophophores, etc.) on both micro- and macroevolutionary timescales. How do we expect variation in morphology, which is shaped by development, growth, or functional constraints, etc., to scale between populations and phylogenies? Can we reconstruct population-level processes from comparative data drawn from a variety of related taxa or from a long stratigraphic sequence? How do we expect inter-taxon patterns of diversity in complex morphologies to be affected by morphological integration, or by different modes of selection, such as directional, stabilizing, or drift? Progress has been made in regards to some of these issues for simple morphological measures, such as size and ratios (Gingerich 1993, 2001; Wagner 1996; Roopnarine 2001, 2003; Polly 2001a). The simulation method presented here should help make further progress towards a morphological synthesis.

The simulations presented here make it clear that certain aspects of microevolutionary process can be detected by comparison of morphologies separated by long stretches of phylogenetic time. The scaling between multivariate morphological distance and the duration of phylogenetic divergence is linear under directional selection, curvilinear under randomly varying selection or drift, and stochastically constant under stabilizing selection. Comparison of the teeth of different populations, species, and genera of shrew demonstrates that the observed scaling between distance and divergence is similar to that under fluctuating selection.

Furthermore, comparison of empirical data to simulation results demonstrates that molar tooth form does not evolve by purely genetic drift. Drift is a function of morphological variance and population size, and is more significant at very small effective population sizes. Even when the mean effective population size is assumed to be only 70 individuals, drift does not produce enough change over even millions of generations to explain the observed morphological divergences in shrew

molar shape. Consequently, the mode of evolution in shrew molar morphology is likely selection that fluctuates from generation to generation in intensity and direction.

The simulations also indicate that morphometric shape is likely to have a strong phylogenetic component that can be utilized for phylogeny reconstruction. When the simulation is given phylogenetic structure, related tip taxa share visible derived similarity. Closely related taxa are also located close to one another in principal components space, locations that can be estimated using principal components analysis of the tip taxa. Maximum-likelihood methods for continuous traits (Felsenstein 1973) are appropriate for estimating trees from such data because they are based on the assumption of Brownian-motion evolution, which is equivalent to the randomly fluctuation selection patterns found here. Previous empirical studies of morphometric-based phylogenetic trees have found that maximum-likelihood trees based on shape data correspond well to the patterns of relationship suggested by independent data (Polly 2003a, 2003b).

The simulation method presented here could be elaborated to incorporate more complicated evolutionary assumptions. The simulations run in this paper were based on **P**, the phenotypic variance-covariance matrix; however, the same method could be applied to **G**, the additive genetic covariance matrix, if it is available. Furthermore, the simulations could be run using more sophisticated quantitative genetic models. The simplest of these improvements would be to allow **P** to evolve, but more complicated covariance models that incorporate developmental interactions, as well as genetic models could also be used (Johnson and Porter 2001; Wolf et al. 2001; Rice 2002; Salazar-Cuidad and Jernvall 2002).

ACKNOWLEDGEMENTS

J. Wójcik of the Mammal Research Centre, Polish Academy of Sciences in Bialowieza, Poland, provided access to the specimens on which the **P** matrix was based. J. Head, J. Jernvall, S. Le Comber, N. MacLeod, and I. Salazar-Ciudad provided valuable comments and criticisms at various stages of development. This research was supported by grant GR3/12996 from the UK Natural Environment Research Council.

REFERENCES

Adrain, J.M. 2001. Systematic paleontology. *Journal of Paleontology*, 75:1055-1057.

Alroy, J. 2003. Global databases will yield reliable measures of global biodiversity. *Paleobiology*, 29:26-29.

Arnold, S.J. 1992. Constraints in phenotypic evolution. *American Naturalist*, 140:S85-S107.

Arnold, S.J., Pfrender, M.E., and Jones, A.G. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica*, 112-113:9-32.

Arnold, S.J., and Phillips, P.C. 1999. Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the Garter snake *Thamnophis elegans*. *Evolution*, 53:1516-1527.

Atchley, W.R., and Hall, B.K. 1991. A model for development and evolution of complex morphological structures. *Biological Reviews*, 66:101-157.

Avise, J.C. 1994. *Molecular Markers, Natural History, and Evolution.* Chapman and Hall, New York.

Avise, J.C. 2000. *Phylogeography: The History and Formation of Species.* Harvard University Press, Cambridge, Massachusetts.

Avise, J.C., Ball, R.M., and Arnold, J. 1988. Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Molecular Biology and Evolution*, 5:331-334.

Badyaev, A., and Hill, G.E. 2000. The evolution of sexual dimorphism in the house finch. I. Population divergence in morphological covariance structure. *Evolution*, 54:1784-1794.

Berg, H.C. 1993. *Random Walks in Biology*. Princeton University Press, Princeton.

Blackith, R.E., and Reyment, R.A. 1971. *Multivariate Morphometrics*. Academic Press, London.

Bookstein, F.L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:567-585.

Brodie, E.D., III. 1993. Homogeneity of the genetic variance-covariance matrix for antipredator traits in two natural populations of the Garter snake, *Thamnophis ordinoides*. *Evolution*, 47:844-854.

Brown, W.M., George, M. Jr., and Wilson, A.C. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences USA*, 76:1967-1971.

Butler, P.M. 1961. Relationships between upper and lower molar patterns, pp. 117-126. In Vanderbroek, G. (ed.), *International Colloquium on the Evolution of Lower and Non-specialized Mammals, Vol. 1*. Paleis der Academiën, Brussels.

Cheetham, A.H., Jackson, J.B.C., and Hayek, L.C. 1993. Quantitative genetics of bryozoan phenotypic evolution. I. Rate tests for random change versus selection in differentiation of living species. *Evolution*, 47:1526-1538.

Cheetham, A.H., Jackson, J.B.C., and Hayek, L.C. 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.

Cheetham, A.H., Jackson, J.B.C., and Hayek, L.C. 1995. Quantitative genetics of bryozoan phenotypic evolution. III. Phenotypic plasticity and the maintenance of genetic variation. *Evolution*, 49:290-296.

Cheverud, J.M. 1995. Morphological integration in the Saddle-back Tamarin (*Saguinus fuscicollis*) cranium. *American Naturalist*, 145:63-89.

Churchfield, S. 1982. Food availability and diet of the Common shrew, *Sorex araneus*, in Britain. *Journal of Animal Ecology*, 51:15-28.

Churchfield, S. 1990. *The Natural History of Shrews*. Christopher Helm, London.

Churchfield, S. 2004. Personal communication.

Clyde, W.C., and Gingerich, P.D. 1994. Rates of evolution in the dentition of Early Eocene *Cantius*: comparison of size and shape. *Paleobiology*, 20:506-522.

Crompton, A.W. 1971. The origin of the tribosphenic molar. *Zoological Journal of the Linnean Society*, 50 (suppl. 1):65-87.

Dehm, R. 1962. Altpleistocäne Säuger von Schernfeld bei Eichstätt in Bayern. Mitteilungen der Bayerischen Staatssammlung für Paläontologie und Historishe Geologie, 2:17-61.

Dryden, I.L., and Mardia, K.V. 1998. *Statistical Shape Analysis*. John Wiley and Sons, New York.

Evans, A.R., and Sanson, G.D. 2003. The tooth of perfection: functional and spatial constraints on mammalian tooth shape. *Biological Journal of the Linnean Society*, 78:173-191.

Falconer, D.S. and Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*, 4th Edition. Longman: Harlow, England.

Felsenstein, J. 1973. Maximum likelihood estimation of evolutionary trees from continuous characters. *American Journal of Human Genetics*, 25:471-492.

Felsenstein, J. 1981. Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. *Evolution*, 35:1229-1242.

Felsenstein, J. 1988. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*, 19:445-471.

Felsenstein, J. 2002. Quantitative characters, phylogenies, and morphometrics, pp. 27-44. In MacLeod, N., and Forey, P. (eds.), *Morphology, Shape, and Phylogenetics*. Taylor & Francis, London.

Foote, M. 1994. Morphological disparity in Ordivician-Devonian crinoids and the early saturation of morphological space. *Paleobiology*, 20:320-344.

Foote, M. 1997. The evolution of morphological diversity. *Annual Reviews of Ecology and Systematics*, 28:129-152.

Foote, M. 1999. Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. *Paleobiology*, 25 (supplement):1-115.

Fumagalli, L., Taberlet, P., Stewart, D.T., Gielly, L., Hausser, J., and Vogel, P. 1999. Molecular phylogeny and evolution of *Sorex* shrews (Soricidae: Insectivora) inferred from mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution,* 11:222-235.

Gingerich, P.D. 1993. Quantification and comparison of evolutionary rates. *American Journal of Science*, 293-A:453-478.

Gingerich, P.D. 2001. Rates of evolution on the time scale of the evolutionary process. *Genetica*, 112-113:127-144.

Gingerich, P.D. and Uhen, M.D. 1998. Likelihood estimation of the time of origin of Cetacea and the time of divergence of Cetacea and Artiodactyla. *Palaeontologia Electronica,* 1(2): 45 pp. [http://palaeo-electronica.org/1998_2/ging_uhen/issue2.htm].

Golub, G.H.,and Van Loan, C.F. 1983. *Matrix Computations*. Johns Hopkins University Press, Baltimore.

Gower, J.C. 1975. Generalized Procrustes analysis. *Psychometrika,* 40:33-51.

Hansen, T.F.,and Martins, E.P. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution*, 50:1404-1417.

Harris, A.H. 1998. Fossil history of shrews in North America, pp. 131-156. In Wójcik, J.M., and Wolsan, M. (eds.), *Evolution of Shrews*, Mammal Research Institute, Polish Academy of Sciences, Bialowieza.

Hewitt, G.M. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68:87-112.

Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature*, 405:907-913.

Humphries, C.J. 2002. Homology, characters, and continuous variables, pp. 8-26. In MacLeod, N., and Forey, P.L. (eds.), *Morphology, Shape, and Phylogeny*. Taylor & Francis, London.

Ibrahim, K.M., Nichols, R.A., and Hewitt, G.M. 1995. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, 77:282-291.

Janis, C.M., Damuth, J., and Theodor, J.M. 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. *Palaeogeography, Palaeoclimatology, and Palaeoecology*, 207:371-398.

Jernvall, J., Hunter, J.P., and Fortelius, M. 1996. Molar tooth diversity, disparity, and ecology in Cenozoic ungulate radiations. *Science*, 274:1489-1492.

Johnson, N.A., and Porter, A.H. 2001. Towards a new synthesis: population genetics and evolutionary developmental biology. *Genetica*, 112-113:45-58.

Jolicoeur, P., and Mosimann, J.E. 1960. Size and shape variation in the Painted Turtle. A principal component analysis. *Growth*, 24:339-354.

Kimura, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Kingsolver, J.G, Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gilbert, P., and Beerli, P. 2001. The strength of phenotypic selection in natural populations. *American Naturalist*, 157: 245-261.

Kinnison, M.T, and Hendry, A.P. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica*, 112-113:145-164.

Klicka, J., and Zink, R.M. 1997. The importance of recent ice ages in speciation: a failed paradigm. *Science*, 277:1666-1669.

Klicka, J., and Zink, R.M. 1998. Pleistocene effects of North American songbird evolution. *Proceedings of the Royal Society of London B*, 266: 695-700.

Klingenberg, C.P., and Leamy, L.J. 2001. Quantitative genetics of geometric shape in the mouse mandible. *Evolution*, 55:2342-2352.

Knowles, L.L. 2004. The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology*, 17:1-10.

Kohn, L.A.P., and Atchley, W.R. 1988. How similar are genetic correlation structures? Data from rats and mice. *Evolution*, 42:467-481.

Kurtén, B., and Anderson, E. 1980. *Pleistocene Mammals of North America*. Columbia University Press, New York.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30:314-334.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution*, 33:402-416.

Lande, R. 1986. The dynamics of peak shifts and the pattern of morphological evolution. *Paleobiology*, 12:343-354.

Lande, R., and Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution*, 37:1210-1226.

Lieberman, B.S., Brett, C.E., and Eldredge, N. 1995. A study of stasis and change in two species lineages from the Middle Devonian of New York State. *Paleobiology*, 21:15-27.

Lofsvold, D. 1986. Quantitative genetics of morphological differentiation in *Peromyscus*. I. Tests of homogeneity of genetic covariance among species and subspecies. *Evolution*, 40:559-573.

Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *American Naturalist*, 136:727-741.

Lynch, M., and Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits.* Sinauer, Sunderland, Massachusetts.

MacLeod, N. 2002. Phylogenetic signals in morphometric data, pp. 100-138. In MacLeod, N., and Forey, P.L. (eds.), *Morphology, Shape, and Phylogeny*. Taylor & Francis, London.

MacLeod, N., and Rose, K.D. 1993. Inferring locomotor behavior in Paleogene mammals via eigenshape analysis. *American Journal of Science*, 293-A:300-355.

Marroig, G., and Cheverud, J.M. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. *Evolution* 55:2576-2600.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D., and Wolpert, L. 1985. Developmental constraints and evolution. *Quarterly Review of Biology*, 60:265-287.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Olson, E.C., and Miller, R.L. 1958. *Morphological Integration*. University of Chicago Press, Chicago.

Padian, K., Lindberg, D., and Polly, P.D. 1994. Cladistics and the fossil record: the uses of history. *Annual Review of Earth and Planetary Sciences*, 22:63-91.

Patton, J.L., and Smith, M.F. 1989. Population structure and the genetic and morphologic divergence among pocket gopher species (genus *Thomomys*), pp. 284-304. In Otte, D., and Endler, J.A. (eds.), *Speciation and Its Consequences*. Sinauer, Sunderland, Massachusetts.

Polly, P.D. 2001a. On morphological clocks and paleophylogeography: Towards a timescale for Sorex hybrid zones. *Genetica*, 112-113:339-357.

Polly, P.D. 2001b. Paleontology and the comparative method: Ancestral node reconstructions versus observed node values. *American Naturalist*, 157:596-609.

Polly, P.D. 2002. Phylogenetic tests for differences in shape and the importance of divergence times: Eldredge's enigma explored, pp. 220-246. In MacLeod, N., and Forey, P.L. (eds.), *Morphology, Shape, and Phylogenetics*. Taylor & Francis, London.

Polly, P.D. 2003a. Paleophylogeography of *Sorex araneus*: molar shape as a morphological marker for fossil shrews. *Mammalia*, 68:233-243.

Polly, P.D. 2003b. Paleophylogeography: the tempo of geographic differentiation in marmots (*Marmota*). *Journal of Mammalogy*, 84:278-294.

Polly, P.D. 2005. Development, geography, and sample size in P matrix evolution: molar-shape change in island populations of *Sorex araneus*. *Evolution and Development*, 7:29-41.

Polly, P.D., Le Comber, S.C., and Burland, T.M. In press. On the occlusal fit of tribosphenic molars: are we underestimating species diversity in the Mesozoic? *Journal of Mammalian Evolution*, 12.

Ratkiewiezc, M., Fedyk, S., Banaszek, A., Gielly, L., Chêtnicki, W., Jadwiszczak, K., and Taberlet, P. 2002. The evolutionary history of two karyotypic groups of the Common shrew, *Sorex araneus*, in Poland. *Heredity*, 88:235-242.

Raup, D.M. 1967. Geometric analysis of shell coiling: coiling in ammonoids. *Journal of Paleontology*, 41:43-65.

Raup, D.M. 1976. Species diversity in the Phanerozoic: a tabulation. *Paleobiology*, 2:279-288.

Raup, D.M., and Gould, S.J. 1974. Stochastic simulation and evolution of morphology – towards a nomothetic paleontology. *Systematic Zoology*, 23:305-322.

Raup, D.M., and Marshall, L.G. 1980. Variation between groups in evolutionary rates: a statistical test of significance. *Paleobiology*, 6:9-23.

Repenning, C.A. 1967. Subfamilies and genera of the Soricidae. *U.S. Geological Survey Professional Paper*, 565:1-74.

Rice, S.H. 2002. A general population genetic theory for the evolution of developmental interactions. *Proceedings of the National Academy of Sciences, USA*, 99:15518-15523.

Roff, D.A. 1997. *Evolutionary Quantitative Genetics*. Chapman & Hall, New York.

Roff, D.A., Mosseau, T., Møller, A.P., de Lope, F., and Saino, N. 2004. Geographic variation in the G matrices of wild population of the Barn swallow. *Heredity*, 93:8-14.

Rohlf, F.J. 1999. Shape statistics: Procrustes superimpositions and tangent spaces. *Journal of Classification* 16:197-223.

Rohlf, F.J., and Slice, D. 1990. Extension of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, 39:40-59.

Roopnarine, P.D. 2001. The description and classification of evolutionary mode: a computational approach. *Paleobiology*, 27:446–465.

Roopnarine, P.D. 2003. Analysis of rates of morphological evolution. *Annual Review of Ecology and Evolution*, 34:605-632.

Rzebik-Kowalska, B. 1998. Fossil history of shrews in Europe, pp. 23-92. In Wójcik, J.M., and Wolsan, M. (eds.), *Evolution of Shrews*, Mammal Research Institute, Polish Academy of Sciences, Bialowieza.

Salzar-Ciudad, I., and Jernvall, J. 2002. A gene network model accounting for development and evolution in mammalian teeth. *Proceedings of the National Academy of Sciences, USA*, 99:8116-8120.

Schmalhausen, I.I. 1949. *Factors of Evolution: The Theory of Stabilizing Selection*. Translated by T. Dobzhansky. Blakiston, Philadelphia.

Semken, H.A., Jr. 1984. Paleoecology of a late Wisconsinan/Holocene micromammal sequence in Peccary Cave, northwestern Arkansas. *Carnegie Museum of Natural History Special Publication*, 8:405-431.

Sepkoski, J.J. 1997. Biodiversity: past, present, and future. *Journal of Paleontology*, 71:533-539.

Sheets, H.D., and Mitchell, C.E. 2001. Uncorrelated change produces the apparent dependency of evolutionary rate on interval. *Paleobiology*, 27: 429-445.

Simpson, G.G. 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.

Simpson, G.G. 1953. *The Major Features of Evolution*. Columbia University Press, New York.

Slice, D.E., Rohlf, F.J., and Bookstein, F.L. 1996. Glossary, pp. 531-551. In Marcus, L.F., Corti, M., Loy, A., Naylor, G., and Slice, D.E. (eds.), *Advances in Morphometrics*. Plenum, New York.

Smith, L.H. 1998. Species-level phenotypic variation in lower Paleozoic trilobites. *Paleobiology*, 24:17-36.

Smith, L.H., and Lieberman, B.S. 1999. Disparity and constraint in olenellid trilobites and the Cambrian radiation. *Paleobiology*, 25:459-470.

Spicer, G.S. 1993. Morphological evolution of the *Drosophila virilis* species group as assessed by rate tests for natural selection on quantitative characters. *Evolution*, 47:1240-1254.

Springer, M.S. 1997. Molecular clocks and the timing of placental and marsupial radiations in relation to the Cretaceous-Tertiary boundary. *Journal of Mammalian Evolution*, 4:285-302.

Steppan, S.J. 1997. Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal components analyses. *Evolution,* 51:571-586.

Storch, G., Qiu, Z., and Zazhigin, V.S. 1998. Fossil history of shrews in Asia, pp. 91-120. In Wójcik, J.M., and Wolsan, M. (eds.), *Evolution of Shrews*, Mammal Research Institute, Polish Academy of Sciences, Bialowieza.

Tatsuoka, M.M., and Lohnes, P.R. 1988. *Multivariate Analysis: Techniques for Educational and Psychological Research.* Macmillan, New York.

Turelli, M., Gillespie, J. H., and Lande, R. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. *Evolution*, 42:1085-1089.

Vrba, E.S., and Eldredge, N. 1984. Individuals, hierarchies, and processes: towards a more complete evolutionary theory. *Paleobiology*, 10:146-171.

Wagner, P.J. 1996. Contrasting the underlying patterns of active trends in morphologic evolution. *Evolution*, 50:990-1007.

Wakeley, J., and Hey, J. 1998. Testing speciation models with DNA sequence data, pp. 157-175. In DeSalle, R., and Schierwater, B. (eds.), *Molecular Approaches to Ecology*. Birkhäuser Verlag, Basel.

Wójcik, J.M. 2004. Personal communication. N_e in Bia³owie¿a populations of *Sorex araneus*.

Wolf, J.B., Frankino, W.A., Agrawal, A.F., Brodie, E.D. III, and Moore, A.J. 2001. Developmental interactions and the constituents of quantitative variation. *Evolution*, 55:232-245.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics*, 16:97-159.

Wright, S. 1968. *Evolution and genetics of populations. Vol. 1. Genetics and Biometric Foundations.* University of Chicago Press, (Chicago?.)

Zelditch, M.L., Bookstein, F.L., and Lundrigan, B.L. 1992. Ontogeny of integrated skull growth in the Cotton Rat, *Sigmodon fulviventer*. *Evolution*, 46:1164-1180.

Zima, J., Lukácová, L., and Macholán, M. 1998. Chromosomal evolution in shrews, pp. 173-218. In Wójcik, J.M., and Wolsan, M. (eds.), *Evolution of Shrews*, Mammal Research Institute, Polish Academy of Sciences, Bialowieza.

APPENDIX

The equivalence between Lande's (1979) formulation of multivariate evolution and the shape-space simulation adopted here can be demonstrated as follows. Let **Z** be a matrix of *n* phenotypic traits measured on *m* individuals, which are distributed normally with some amount of covariance. The vector of trait means is $\bar{z}$. Let **P** be the *n* by *n* phenotypic covariance matrix of *z*, and let ***â*** be a vector of selection differentials of length *n* postmultiplied by matrix **H** representing the heritability of **P**. Each selection differential is the standardized difference between trait means in selected and unselected adults (Lande 1979; Falconer and Mackay 1996).

Lande's equation for change in the multivariate mean in response to selection over one generation can then be written as follows:

$$\Delta\bar{z} = \beta HP \quad \text{(A1)}$$

where $\Delta\bar{z}$ is the change in means of traits **Z** after one generation of selection (Lande 1979). Lande's original formula used **G**, the additive-genetic covariance matrix, instead of **P**, but **G** is equal to the heritable portion of the phenotypic covariance matrix, or **HP**.

The morphological shape-space in which the simulation is performed is described by matrices obtained by the singular-value decomposition (SVD) of **P**,

$$P = U^T W V \quad \text{(A2)}$$

where **U** is the left singular vector matrix representing the weights of the original variables on the principal component vectors (equivalent to the eigenvectors, **Table 2**), **W** is the diagonal matrix of the singular values representing the variances of the data on the principal components vectors (equivalent to the eigenvalues, **Table 2**), **V** is the right singular vector matrix, and the superscript ^{T} represents the transpose of the matrix (Golub and Van Loan 1983). **U** functions as a *j* by *n* rotation matrix used to transform data from the original coordinate system to the principal components axes (or back again), where *j* is the matrix rank of **P**. When **P** is not singular then *j* = *i*, but when **P** is singular, as with Procrustes superimposed shape data, *j* is smaller than *i*. **W** is a *j* length diagonal matrix that can be thought of as **P** transformed so that trait covariances are zero. Note that for real, positive definite symmetric matrices **U** and **V** are equal. The projections of the original data onto the principal components (the scores) are given by

$$Z' = ZU^T \quad \text{(A3)}$$

where Z' are the scores (Jolicoeur and Mosimann 1960). The total variances in **Z** and Z' are equal, but covariances are absent in Z' and the variances of Z' are given by **W**.

Now, let $\bar{z}$ at time *0* be $\bar{z}_0$ = {0, 0, … *i*}. The projection of $\bar{z}_0$ onto the principal components axes is $\bar{z}_0 U^T$ which equals {0, 0, … *j*}. The vector of mean morphologies in the next generation, $\bar{z}_1$, equals $\Delta\bar{z}$ or β**H****P**. The projection of $\bar{z}_1$ into the principal components space is

$$\bar{z}_1' = \bar{z}_1 U^T \quad \text{(A4)}$$

which, by substitution, equals

$$\begin{aligned}\Delta\bar{z}' &= \beta HPU^T \\ &= \beta HU^T WUU^T \\ &= \beta HU^T W \\ &= (\beta H)' W\end{aligned} \quad \text{(A5)}$$

Importantly, solving for $\bar{z}_1$ in Eq. A4 reveals that the transformation back into the original coordinate space only requires multiplication by **U**, such that

$$\Delta\bar{z} = (\beta H)' WU \quad \text{(A6)}$$

Equations A5 and A6 show that simulations can be carried out in principal components space using only a vector of selection differentials and **W**, which can be treated as a vector of scaling weights. A simulation in reduced-dimension

shape-space is thus computationally more efficient than one in the original space because $(\beta H)'$ is of length *j* and requires fewer operations than ***â*H** of length *n* and because multiplication by vector **W** takes fewer calculations than multiplication by covariance matrix **P**. At the end of the simulation, the results can easily be transformed back to the original coordinate space by multiplying them by **U**.

Simulation in principal components space also overcomes complications introduced by the singularity of **P** when it is based on Procrustes superimposed data. Procrustes fitting removes rotation, translation, and scaling from the data, resulting in a covariance matrix of reduced rank (Rohlf 1999; Dryden and Mardia 1998). The singularity can produce seemingly odd results when a uniform vector of selection differentials is applied to Eq. A1. For example if ***â*H** = {1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1} were applied to **P** from the Bialowieza sample, the resulting $\Delta\bar{z}$ = {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}. This result is counterintuitive because we seem to have applied uniform positive selection to the system but have obtained no change in phenotype. The colinearity of the matrix has the effect of cancelling the effects of the uniform selection vector. When uniform positive selection is applied in principal components space as $(\beta H)'$ = {1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1} we get a pleasing $\Delta\bar{z}'$ that moves the mean phenotypic in a positive direction along each principal component. The same change could be produced in the original coordinate space, but only by using the counterintuitive vector of selection differentials ***â*H** = {-0.121, -0.012, 0.268, 0.883, -0.929, -1.114, -0.199, -0.432, 2.464, 0.797, 0.219, 0.824, -0.699, -0.835, -1.58, 0.343, 0.583, -0.454}. This situation makes an intuitive simulation of positive directional selection exceedingly difficult unless it is performed in the reduced dimension shape-space.