Phylogenies and comparative data, a microevolutionary perspective

E. P. MARTINS

Department of Biology, University of Oregon, Eugene, Oregon 97403, U.S.A.

SUMMARY

As species evolve along a phylogenetic tree, their phenotypes diverge. We expect closely related species to retain some phenotypic similarities owing to their shared evolutionary histories. The degree of similarity depends both on the phylogeny and on the detailed evolutionary changes that accumulate each generation. In this study, I review a general framework that can be used to translate between macroevolutionary patterns and the underlying microevolutionary process by comparing the observed relationships among measured species phenotypes and the expected relationship structure due to the phylogeny and underlying models of phenotypic evolution. I then show how the framework can be used to compare methods used (1) to reconstruct phylogenies, (2) to correct comparative data for phylogenetic non-independence, and (3) to infer details of the microevolutionary process from interspecific data and a phylogeny. Use of this framework and a microevolutionary perspective on the analysis of interspecific data opens up new fields of inquiry and many new uses for phylogenies and comparative data.

1. INTRODUCTION

Usually, interspecific or 'comparative' data are not independent of one another because the species from which the data are measured have been evolving together for some period of evolutionary time (see, for example: Felsenstein 1985; Harvey & Pagel 1991). The exact amount of non-independence depends both on the phylogeny by which the species are related and on the underlying processes of phenotypic evolution working at each generation.

Traditionally, systematists have used this non-independence of comparative data to infer the branching patterns of speciation (i.e. the phylogeny) underlying extant organisms (see, for reviews: Felsenstein 1988; Swofford & Olsen 1990). Several phenotypic traits are chosen that fit the assumptions of some numerical algorithm (e.g. traits that are believed to have evolved neutrally or parsimoniously), and the relationships between interspecific measurements of these traits are used to infer the historical relationships between the measured species. Throughout this procedure, the emphasis is on the evolution of species (or other taxonomic group).

In contrast, the 'comparative method' is a family of techniques in which interspecific measurements are used to infer something about the biology of particular traits. Independence of data points is one of the primary assumptions of most parametric statistical procedures; so, when ordinary statistics are used to analyse comparative data, this assumption is regularly violated. In recent years, a number of techniques have been proposed to correct this problem by incorporating phylogenetic or taxonomic information into the analysis (see, for reviews: Harvey & Pagel 1991; Miles & Dunham 1994; Maddison 1994; Martins & Hansen 1995). Unlike phylogeny reconstruction, in the statistical correction of comparative data, the emphasis is on the organismal traits themselves rather than on the species that exhibit those traits. Traits are chosen because of some particular hypothesis, rather than because they fit a set of predefined assumptions, and generally the phylogeny and models of phenotypic evolution are assumed to be known.

As new phylogenetic comparative methods were developed, researchers also began to consider the many other evolutionary questions that can be answered given a set of interspecific measurements and a phylogeny. Using comparative data, we can now infer the ancestral states of phenotypes, the strength and type of microevolutionary forces acting on characters, the relationships between evolutionary changes in two or more traits, and the degree of phylogenetic inertia in a character.

Thus, there have been three primary uses of interspecific data in evolutionary biology: (1) to reconstruct phylogenies, (2) to correct a problem in the statistical analysis of comparative data, and (3) to infer the detailed evolutionary history of particular characters. In this paper, I discuss how a microevolutionary perspective can be used to link these three processes, and thereby clarify the issues underlying all three. Using the general framework proposed in Hansen & Martins (1995), I discuss what can and cannot be inferred from comparative data and a phylogeny, and how that framework can be used to evaluate and compare proposed methods. In many cases, statistical techniques to estimate the desired parameters have not

Phil. Trans. R. Soc. Lond. B (1995) **349**, 85–91
Printed in Great Britain

yet been developed and/or phylogenetic information may not be adequate to answer the questions. At the risk of being overly optimistic, this paper strives only to discuss what *may* be possible given appropriate information.

2. GENERAL FRAMEWORK

The following is an applied version of the general framework proposed by Hansen & Martins (1995). If we measure the phenotypes or genotypes of several species in a clade, we can describe the relationship between the measurements of each pair of species as a symmetrical $N \times N$ matrix, where N is the number of species in the clade. I shall refer to this matrix as the *observed relation matrix* (ORM; e.g. table 1). The elements of the ORM may be phenetic distances, estimates of genetic variance and covariance, or any other function of the measured data. The ORM can be for a single character or multiple characters (in which case the ORM would be multidimensional).

We can obtain a complementary matrix of the expected pattern of relationship among measured species by considering the evolutionary processes that led to those relationships. At each generation, natural selection and random genetic drift interact with heredity and environmental fluctuations to form the observed phenotype. Most of these forces are likely to be stochastic, such that the phenotype at each generation can be viewed as a probabilistic phenomenon with some mean expectation and variance about that expectation. During multiple generations and speciation events, evolution unfolds along the branches of a phylogenetic tree resulting in phenotypes of extant species. Because the microevolutionary forces involved are stochastic, evolution along a phylogeny might have occurred in any one of an infinite number of possible ways. The evolutionary pathway that actually occurred can be viewed as a single sample from this statistical population of possible evolutionary scenarios.

The single result of the 'true' evolutionary scenario (i.e. the phenotypes of extant species) can be described as an ORM. Similarly, the endpoints of each of the other possible scenarios in the statistical population can also be described as relationship matrices. These relationship matrices will share some statistical properties (e.g. means, variances, covariances) determined by the microevolutionary forces acting throughout evolution. Given information about those microevolutionary forces, we can infer the expected relationships between all possible pairs of species, and describe them as an *expected relation matrix* (ERM; table 2).

Table 1. *An example with use of hypothetical comparative data*

species	mean phenotypes
A	2.0
B	3.0
C	1.5
D	1.0
E	4.5
F	3.5

sample observed relation matrix (ORM)

	A	B	C	D	E	F
A	0.34	0.90	0.30	0.30	0.28	0.34
B	0.90	0.35	0.31	0.27	0.29	0.32
C	0.30	0.31	0.31	0.85	0.31	0.34
D	0.30	0.27	0.85	0.30	0.32	0.29
E	0.28	0.29	0.31	0.32	0.29	0.88
F	0.34	0.32	0.34	0.29	0.88	0.30

Table 2. *Examples of expected relation matrices for phylogenies in figure 1.*

(Phylogeny A, figuure 1*a*. Under a Brownian motion model of gradual phenotypic evolution, $V_A = \sigma t$, where σ is the variance of evolutionary changes and t is the time from the root of the tree to the present. Under a speciational or punctuational version of the same model, $V_A = \sigma t$, still, but t is the number of speciation events from the root of the tree to the present (in this case, one).

Phylogeny B, figure 1*b* . Under a Brownian motion model of gradual phenotypic evolution, $V_A = \sigma t$, where σ is the variance of evolutionary changes and t is the time from the root of the tree to the present. Similarly, $C_B = \sigma t_A$, where t_A is the time that each pair of sister species evolved together from the root of the tree to the point of diversification into different species. Under a speciational or punctuational version of the same model, $V_A = \sigma t$ and $C_B = \sigma t_A$ still, but t is measured in units of the number of speciation events. In this case, t is the number of speciation events occurring from the root of the tree to the present (i.e. two) and t_A is the number of speciation events occurring from the root of the tree to the point of diversification of the pairs of sister species (one).)

expected relation matrices (ERMs) for:

phylogeny A

	A	B	C	D	E	F
A	V_A	0	0	0	0	0
B	0	V_A	0	0	0	0
C	0	0	V_A	0	0	0
D	0	0	0	V_A	0	0
E	0	0	0	0	V_A	0
F	0	0	0	0	0	V_A

phylogeny B

	A	B	C	D	E	F
A	V_B	C_B	0	0	0	0
B	C_B	V_B	0	0	0	0
C	0	0	V_B	C_B	0	0
D	0	0	C_B	V_B	0	0
E	0	0	0	0	V_B	C_B
F	0	0	0	0	C_B	V_B

3. EFFECTS OF VARYING PHYLOGENY AND MODEL OF CHANGE – SOME EXAMPLES

Imagine a star phylogeny in which all species in clade A diverged instantaneously from a single ancestor and have been evolving by the same processes independently of each other ever since (figure 1*a*). The expected relationship (e.g. covariance) between the phenotypes of each pair of species on this phylogeny will be the same as the expected covariance between all other pairs of species on this phylogeny, because each

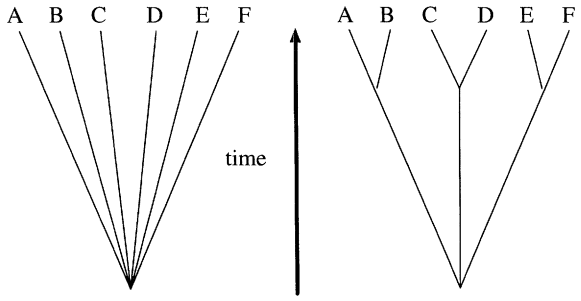

Figure 1. Two phylogenies for use with the examples.

species has been evolving independently of every other species for exactly the same amount of time and under the same sort of microevolutionary processes. Thus, the off-diagonal elements of the ERM will be identical (see table 2). The variance of each species also depends on the stochastic nature of evolution such that if each species has been evolving for the same amount of total time and under the same sorts of microevolutionary processes, the expected variance of all species will be identical. Thus, the ERM in this situation is a matrix with diagonal elements V_A, and off-diagonal elements, C_A.

Now imagine a phylogeny in which two major speciation events occurred (figure 1*b*). In this example, the ancestral lineage divided once into three taxa, each of which later divided again into two sister species, resulting in a total of six taxa (figure 1*b*). Each species is expected to be as similar to its sister species as each other species in the clade is to its sister species, since all sister species have been evolving together or independently of one another for the same amount of time and under the same sorts of microevolutionary processes. All sister species will thus share a single value of expected covariance (C_{B1}) (see table 2). For the same reasons, each species is also expected to be as similar to its non-sister taxa as is every other species in the clade (C_{B2}). Again, the expected variance of each species will be the same as the expected variance of all other species on the phylogeny (V_B) because they have all been evolving for the same total length of time.

The ERMs of clades A and B will also depend on the details of the microevolutionary process underlying phenotypic change (e.g. table 2). For example, if phenotypic evolution is due solely to random genetic drift in a simple model of neutral evolution, similarity among species will depend primarily on the time that the species have been evolving together or independently of one another. In this case, we should expect $V_A = V_B$, because the species in clade A have been evolving for the same total amount of time as the species in clade B. On the other hand, we should expect $C_{B2} < C_{B1}$, because in clade B each pair of sister species has evolved together as a single common ancestor for more time than the three sister species ancestors evolved together as the single ancestor of the entire clade. Similarly, we expect $C_A = C_{B2} = 0$ because the species in clade A and the sister species pairs in clade B have been evolving independently of one another from the root of their phylogenies.

Two common models of neutral evolution assume that (1) the total amount of evolutionary change is proportional to time with some change possible at each generation (i.e. a model of 'gradual' evolution) and (2) the total amount of evolutionary change is proportional to the number of speciation events, with change occurring only at branching points on a phylogeny (i.e. a model of 'punctuated' or 'speciational' evolution). In simple Brownian motion forms of both models (for example, as described by Edwards & Cavalli-Sforza (1964), Felsenstein (1973, 1981) and Lynch & Hill (1986)), variances and covariances of species mean phenotypes will equal σt, where σ is the infinitesimal variance of the stochastic process (i.e. the variance of evolutionary changes occurring at each generation or speciation event) and t is a measure of time that the species have been evolving in total (for estimates of the variance) or with other species as a single common ancestor (for estimates of covariance). Measures of time are in number of generations for a true time-based model, or in number of speciation events for a punctuational type of model. Table 2 shows how differences between these types of models and phylogenies can affect the ERM.

4. PHYLOGENY RECONSTRUCTION

Now imagine that we have actually measured six species in an existing clade (see, for example, table 1), and that we wish to infer the phylogeny underlying the evolution of their phenotypes. We can estimate the ORM empirically, for example, by calculating the phenetic distances between all species in the clade. Given a known model of the microevolutionary process, we can also calculate the ERM predicted for any phylogeny containing these six species (see, for example, table 2). If we calculate the ERM predicted for many different phylogenies, then we might define the 'best' phylogeny as one that gives us the closest match between empirical measurements (ORM) and theoretical expectations (ERM). For example, given the ORM in table 1, and the assumption that the character used to calculate this ORM has been evolving purely by random genetic drift, the phylogeny in figure 1*b* would be judged as a better phylogenetic hypothesis than the one in figure 1*a* because its ERM is more similar in form to the ORM.

In very broad terms, most phylogeny reconstruction methods can be represented as a similar comparison of ORM and ERM. Different methods use different criteria for determining whether the ORM and ERM are similar. For example 'distance' methods generally use some sort of least squares criterion, such that the sum of squared differences between ORM and ERM is minimized. Maximum likelihood approaches derive a probability statement from the ERM and then maximize the probability that the ORM was sampled from a population with the characteristics described by the ERM. Other approaches (e.g. parsimony) do not determine an ERM explicitly, but simply prefer an ORM with certain characteristics. For example, parsimony approaches minimize the number of evolutionary changes occurring along the phylogeny. An ERM that reflects this would have small absolute values of both variances and covariances and an internal structure that clearly reflects the phylogeny.

Once ORM and ERM have been determined, there are still many practical difficulties involved in reconstructing phylogenies. For example, there are 945 possible phylogenies for six species and the number of possible trees increases rapidly with the number of species (Felsenstein 1978). Therefore, it is not usually possible to try all of the possible phylogenies, and numerous algorithms for searching among possible phylogenies have been developed. It is also not clear which measure of 'match' between ORM and ERM is the most appropriate. Computer simulation studies suggest that most of the existing approaches give reasonable answers (see, for example, Kuhner & Felsenstein 1994), and still other methods may be possible. Finally and probably most importantly, the best model of the microevolutionary process is not known and is contentiously argued by systematists. The neutral Brownian motion model of molecular evolution used by Felsenstein (1973, 1981) in his maximum likelihood method, the parsimony model of phenotypic evolution developed by Farris (1970) and other authors and the 'distance method' models all have proponents and detractors (see, for reviews: Felsenstein 1988; Swofford & Olson 1990).

5. 'CORRECTING' FOR PHYLOGENY

Most standard parametric statistics (e.g. regression, ANOVA) make three assumptions when estimating parameters and conducting hypothesis tests. These are that the error terms in general linear models underlying the statistical procedure (1) are statistically independent of one another, (2) have homoscedastic (the same or similar) variances and (3) are normally distributed. From an evolutionary perspective, these assumptions can be translated into assumptions regarding the form of the ORM and thus correspondingly the ERM. If the comparative data are statistically independent, they will be uncorrelated with each other such that the off-diagonal elements of the ORM all equal zero ($C_A = C_{B1} = C_{B2} = 0$ in the examples of table 2). Similarly, the requirement that the error term be homoscedastic translates into a requirement that the diagonal elements of the ORM be equal to one another (all V_A are equal; all V_B are equal). The two assumptions together are equivalent to requiring that the ORM be of the form $\sigma^2\boldsymbol{I}$ where σ^2 is the variance of the transformed contrasts and $\boldsymbol{I}$ is the identity matrix (with numbers 1 the diagonal, and numbers 0 elsewhere). In evolutionary terms, only comparative data measured from species related to one another by star phylogenies (figure 1*a*) or with characters evolving under certain microevolutionary models (e.g. strong stabilizing selection) will have ORMs of this type. The final requirement is that the error terms (and thus, usually, the measured data) be normally distributed. Again, only certain microevolutionary processes will lead to normally distributed species phenotypes with an ORM of the above type.

Different comparative methods propose different ways of ensuring that these assumptions will be met. Most, however, can be viewed as a transformation of the raw species data into statistics that meet these assumptions. Many comparative methods use a known phylogeny and assumed model of phenotypic change to transform an existing ORM into one of the form found in the first ERM of table 2. For example, Felsenstein's (1985) method uses an assumption that the characters have been evolving as if by Brownian motion to generate his 'standardized, independent contrasts'. Brownian motion is a mathematical model commonly used in population genetics to describe evolution under random genetic drift. Felsenstein's procedure might be described as use of a model $\boldsymbol{y}' = \boldsymbol{C}\boldsymbol{y}$ to transform raw species data ($\boldsymbol{y}$) into a set of contrasts ($\boldsymbol{y}'$) that have zero covariance (the off-diagonals of their ORM equal zero) and the same variance (the diagonals of the ORM are all equal) by using the information contained in the ERM for that phylogeny and the Brownian motion model of change (combined in a complex way to form the transformation matrix, $\boldsymbol{C}$). For example, the ERMs described in table 2 can be used to transform data measured from species related to each other by the phylogenies in figure 1 such that the resulting ORMs are of the form $\sigma^2\boldsymbol{I}$. Under Brownian motion, the species mean phenotypes will be normally distributed. Thus, if the Brownian motion model is appropriate for the measured data, Felsenstein's contrasts are guaranteed to fit the three primary assumptions of most parametric statistics.

Other comparative methods also solve the problem of non-independence by using ERMs to transform the data such that the ORM of the transformed data is of the form $\sigma^2\boldsymbol{I}$. For example, Cheverud *et al.* (1985) and Gittleman & Kot (1990) proposed a spatial autocorrelation method for incorporating phylogenetic information into the analysis of interspecific data. Using a first order autoregressive model, they partition variation in the comparative data into a phylogenetic component ($\rho\boldsymbol{W}\boldsymbol{y}$, where ρ is an estimated autocorrelation coefficient, $\boldsymbol{W}$ is a relationship matrix similar to the ERMs described in this paper and $\boldsymbol{y}$ is a vector of the measured data) and a non-phylogenetic component ($\epsilon = \boldsymbol{y} - \rho\boldsymbol{W}\boldsymbol{y}$). The model is again, a way of transforming the raw data ($\boldsymbol{y}$) into independent statistics ($\boldsymbol{y}' = \epsilon = \boldsymbol{y} - \rho\boldsymbol{W}\boldsymbol{y}$) that are expected to fit the assumptions of most standard statistical procedures.

Whereas Felsenstein's (1985) model explicitly assumes a Brownian motion model of phenotypic evolution, the ρ coefficient in the model of Cheverud *et al.* (1985) offers some statistical flexibility as to the actual microevolutionary process. Although the authors of this method do not discuss the microevolutionary processes underlying the assumptions of their model, if we consider the method from the evolutionary framework described above, there is still an ERM underlying their transformation of the raw comparative data into the desired form, as there is for all statistical models of comparative data. For example, according to the spatial autoregressive model, the variance of any species mean phenotype (y_i) will be given by $\mathrm{Var}(\rho\Sigma[\boldsymbol{w}_i\boldsymbol{y}] + \epsilon_i)$ where $\boldsymbol{w}_i$ is the row of the $\boldsymbol{W}$ relationship matrix corresponding to species i, and ϵ_i is one of a set of independent, homoscedastic, normally distributed variables. The equation for the covariance between species phenotypes is similar. Thus, from an

evolutionary perspective there is a hidden assumption underlying the use of this spatial model that, whatever microevolutionary processes have been acting in the clade, they have resulted in species data with variances and covariances of the above form. Unfortunately, although several stochastic models of the micro-evolutionary process have been considered in modern evolutionary biology (with and without selection of various types), none of the commonly used models are expected to produce species data with the above form (Martins & Hansen 1995). Similar consideration of other methods for the analysis of comparative data can be useful in comparing methods in terms of their underlying assumptions.

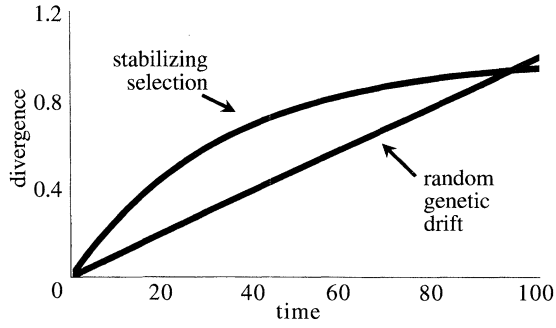

Figure 2. Relation between phenotypic divergence among species and time expected under a microevolutionary model of neutral evolution via random genetic drift alone (straight line) and a model of evolution under stabilizing selection and random genetic drift (curved line).

6. OTHER EVOLUTIONARY QUESTIONS

With the availability of so many approaches to the problem of incorporating phylogenetic information into comparative analyses, researchers have also begun to explore a number of other interesting evolutionary questions that can be answered by combining comparative data with phylogenies. For example, comparative analyses can be used to estimate the degree of 'phylogenetic effect' in a character, the magnitude of the relationship between evolutionary changes in two characters, or the rate of phenotypic evolution of different characters. They can be used to infer the ancestral states of characters and to test whether selection has been acting on a character or group of organisms, and to estimate the strength of that selection. Although statistical methods to conduct these analyses have not all been fully developed, new methods are being proposed every year, and I shall concentrate on discussion of what techniques are theoretically possible, rather than which have already been implemented.

To answer the above questions requires reference to an explicit microevolutionary framework and a known phylogeny. As with phylogeny reconstruction and the correction of statistical problems, the comparison of ORM and ERM provides that framework. In essence, given a known phylogeny and a model of the phenotypic evolutionary process, we can estimate the ERM that corresponds to those assumptions. Given an explicit ERM, we can compare the underlying model of the microevolutionary process with the actual patterns observed in a set of comparative data (the ORM) and use the relations between these two matrices to answer the above questions.

For example, as pointed out originally by Cheverud *et al.* (1985), a measure of the statistical fit of a phylogenetic transformation model to a set of inter-specific data can be used as a reasonable estimate of the degree of phylogenetic 'effect' or 'inertia' inherent in a character. With any method based on standard regression techniques (see, for example: Cheverud *et al.* 1985; Grafen 1989; Lynch 1991), a coefficient of determination (r^2) can be used as a reasonable estimate of the fit of the model to the data. This statistic summarizes the correspondence between the patterns found in the measured comparative data (ORM) and the statistical or evolutionary model (ERM), and can thus serve as a reasonable estimate of the 'phylogenetic effect' of a character. In theory, other measures of similarity can be developed for other methods such as Felsenstein's maximum likelihood approach and parsimony.

Another important question is whether a character has been evolving purely by random genetic drift or whether selection has played an important role. In Martins (1994), I developed a generalized least squares (GLS) procedure to estimate the rate of phenotypic evolution that can also be used as a test of whether the Brownian motion model of phenotypic evolution via random genetic drift provides an adequate fit to a set of comparative data, or whether a model of stabilizing selection (such as that developed by Lande (1976, 1979)) would be more appropriate. This method consists of calculating the divergence between pairs of species on a phylogeny (the ORM) and relating this divergence to the time that the species have been evolving independently of one another. Under a Brownian motion model of evolution, the relation between divergence and time is expected to be linear, with pairs of species that have been separated for long periods of time also exhibiting greater phenotypic differentiation (this is the ERM; figure 2). Under a model of stabilizing selection, the relation between divergence and time is expected to be exponential rather than linear (figure 2). Given a known phylogeny, either or both models can be fitted to a set of comparative data, and the fit of the models to the data can be assessed by using GLS regression procedures. The relative appropriateness of the two competing models can also be compared by using likelihood ratio tests, and estimates of the internal parameters of the models (e.g. the rate of phenotypic evolution, the strength of stabilizing selection) can be obtained. Again, in theory, similar methods could be developed to compare the fit of any microevolutionary model for which an ERM can be defined.

The relation between ORM and ERM can also be used to develop estimators for the ancestral states of a character, or the relation between evolutionary changes in two characters, or to transform the raw comparative data into phylogenetically relevant units that can then be used in other statistical procedures (as discussed in the previous section). Probably, the most

useful ERMs will be obtained from explicit microevolutionary models. For example, in Hansen & Martins (1995) we describe a general model of phenotypic evolution that assumes only that (1) phenotypic evolution can be described as a Markovian stochastic process, (2) the process unfolds along a branching phylogenetic tree and (3) species evolve independently after a split (i.e. speciation event) on the phylogeny. We further describe the macroevolutionary pattern of relationships among species expected under this general model (the ERM). Most of the mathematical models of phenotypic evolution that have been considered in population and quantitative genetics are special cases of this general framework. Thus, we also show how several population genetic models (e.g. evolution via random genetic drift, with or without stabilizing selection, directional selection or environmental fluctuations) can be described in terms of the expected variance–covariance structure (the ERM) of the resulting extant species phenotypes. As mentioned above, ERMs can also be obtained from different statistical procedures (e.g. spatial autoregression) or numerical algorithms (e.g. parsimony). By developing inferential methods based on different ERMs we can create a link between the macroevolutionary pattern observed in comparative data and the microevolutionary processes underlying change at each generation. We can thus create techniques to estimate parameters such as the strength of selection, rates of mutation and even basic genetic variances and covariances.

7. DISCUSSION

Evolutionary biology is by nature an inferential science. We do empirical studies to monitor current evolutionary forces in the field and then infer that the same forces were prevalent in earlier times. Artificial selection experiments show us how microevolutionary forces can act on existing phenotypes and suggest what the result would be if those forces continued over geological time. Even palaeontologists are often forced to infer the former existence and phenotypes of species from the tiniest traces of evidence in the fossil record. Rarely does evolution occur sufficiently quickly or within our view, so that we can observe and measure it.

One of the few types of information that we have about how evolution actually occurred can be measured as the phenotypic and genetic diversity of existing species. By assuming that extant species are the result of long-term evolution along phylogenetic trees, with shared ancestors reaching back to the very beginning of life, we can work backwards and infer the history of those species and their phenotypes. We can observe existing species to measure their phenotypes, their properties of inheritance and the types of environment in which they live. Using this information and some assumptions about the type and magnitude of microevolutionary processes (e.g. selection, drift) that were active in the past, we can develop hypotheses about the patterns of species diversification and division that led to the patterns of phenotypic and genetic similarity observed in extant species. Similarly, if we assume that certain microevolutionary processes were acting, we can work backwards and infer the ancestral states of particular characters, how quickly they evolved and whether or not evolutionary changes in those characters have been subjected to various types of constraints through evolutionary time. Although there are many problems that still need to be addressed regarding inferences made inappropriately from methods that have already been developed (see, for example, Leroi *et al.* 1994) and although some authors (for example, Harvey & Pagel 1991) may despair at ever having sufficiently good phylogenies or methods that are sufficiently assumption-free, the phylogenetic comparative method remains one of the most powerful techniques in modern evolutionary biology.

I thank Thomas Hansen, Paul Harvey and Joe Felsenstein for the many discussions that led to the ideas in this paper. This work was supported by a grant from the National Science Foundation (no. DEB-9406964).

REFERENCES

Cheverud, J. M., Dow, M. M. & Leutenegger, W. 1985 The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weights among primates. *Evolution* **39**, 1335–1351.

Edwards, A. F. W. & Cavalli-Sforza, L. L. 1964 Reconstruction of evolutionary trees. In *Phenetic and phylogenetic classification* (ed. W. H. Heywood & J. McNeill), pp. 67–76. London: Systematics Association publication no. 6.

Farris, J. S. 1970 Methods for computing Wagner trees. *Syst. Zool.* **19**, 83–92.

Felsenstein, J. 1978 The number of evolutionary trees. *Syst. Zool.* **27**, 27–33.

Felsenstein, J. 1973 Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am. J. hum. Genet.* **25**, 471–492.

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

Felsenstein, J. 1985 Phylogenies and the comparative method. *Am. Nat.* **125**, 1–15.

Felsenstein, J. 1988 Phylogenies from molecular sequences: inferences and reliability. *Rev. Genet.* **22**, 521–565.

Gittleman, J. L. & Kot, M. 1990 Adaptation: statistics and a null model for estimating phylogenetic effects. *Syst. Zool.* **39**, 227–241.

Grafen, A. 1989 The phylogenetic regression. *Phil. Trans. R. Soc. Lond.* B **326**, 119–157.

Hansen, T. F. & Martins, E. P. 1995 Translating between microevolutionary process and macroevolutionary patterns: a general model of the correlation structure of interspecific data. (Submitted to *Evolution*.)

Harvey, P. H. & Pagel, M. D. 1991 *The comparative method in evolutionary biology*. Oxford: Oxford University Press.

Kuhner, M. K. & Felsenstein, J. 1994 A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Molec. Biol. Evol.* **11**, 459–468.

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.

Leroi, A. M., Rose, M. R. & Lauder, G. V. 1994 What does the comparative method reveal about adaptation? *Am. Nat.* **143**, 381–402.

Lynch, M. 1991 Methods for the analysis of comparative data in evolutionary ecology. *Evolution* **45**, 1065–1080.

Lynch, M. & Hill, W. G. 1986 Phenotypic evolution by neutral mutation. *Evolution* **40**, 915–935.

Maddison, D. R. 1994 Phylogenetic methods for inferring the evolutionary history and processes of change in discretely valued characters. *A. Rev. Ent.* **39**, 267–292.

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

Martins & Hansen 1995 Phylogenetic comparative methods: the statistical analysis of interspecific data. In *Phylogenies and the comparative method in animal behavior* (ed. E. P. Martins). Oxford: Oxford University Press. (In the press).

Miles, D. & Dunham, A. 1993 *A. Rev. Ecol. Syst.* **24**, 587–619.

Swofford, D. L. & Olsen, G. J. 1990 Phylogeny reconstruction. In *Molecular systematics* (ed. D. M. Hillis & C. Moritz), pp. 411–501. Sunderland, Massachusetts: Sinauer.