Journal of Classification 18: 109-127 (2001)
DOI: 10.1007/s00357-001-0007-0

A Generalized Permutation Model for the Analysis of Cross-Species Data

François-Joseph Lapointe
Université de Montréal

Theodore Garland, Jr.
University of Wisconsin

Abstract: Many fields of biology employ cross-species comparisons. However, because species descend with modification from common ancestors, and rates of evolution may vary among branches of an evolutionary tree, problems of nonindependence and nonidentical distributions may occur in comparative data sets. Several phylogenetically based statistical methods have been developed to deal with these issues, but two are most commonly used. Independent contrasts attempts to transform the data to meet the i.i.d. assumption of conventional statistical methods. Monte Carlo computer simulations attempt to produce phylogenetically informed null distributions of test statistics. A disadvantage of the former is its ultimate reliance on conventional distributional assumptions, whereas the latter may require excessive information on biological parameters that are rarely known. We propose a phylogenetic permutation method that is akin to the simulation approach but requires less biological input information. We show that the conventional, equally likely (EL) randomization model is a special case of our phylogenetic permutations (PP). An application of the method is presented to test the correlation between two traits with cross-species data.

Keywords: Autocorrelated data; Comparative method; Cross-species data; Nonindependent observations; Permutation test; Phylogenetic tree.

This work was partially supported by a NSERC grant OGP0155251 to F.-J. L. and by NSF grants IBN-9157268 (PYI), DEB-9509343, and DEB-9980967 to T.G. We would like to thank two anonymous referees and the Editor for their useful remarks. We are also grateful to Peter E. Midford for computer programming and for many helpful discussions.

Authors' Addresses: François-Joseph Lapointe, Département de sciences biologiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec, H3C 3J7, Canada; email: Francois-Joseph.Lapointe@umontreal.ca. Theodore Garland, Jr., Department of Zoology, 430 Lincoln Drive, University of Wisconsin, Madison, Wisconsin, 52706, USA; email: tgarland@facstaff.wisc.edu.

1. Introduction

Cross-species comparisons have a long and productive history in ecology, biogeography and phylogenetic studies (Brooks and McLennan 1991; Harvey and Pagel 1991; Eggleton and Vane-Wright 1994; Martins 1996). However, the comparative analysis of observations gathered from multiple species presents a problem for conventional statistical methods by implicitly assuming that these observations are independent and identically distributed (i. i. d.) when computing a statistical test. These assumptions are rarely justified because comparative data are always affected by history (their genealogy, or phylogeny), thus making the observations nonindependent. Ecologists working with spatial data have been dealing with an analogous problem for decades. In the case of ecological data, the nonindependence among observations is the result of spatial autocorrelation, whereas phylogenetic autocorrelation is involved with cross-species data. In the phylogenetic case, the usual situation is to find positive autocorrelation: closely related species are more similar than distant species with respect to a certain trait. Regardless of the source of autocorrelation, the statistical problems are similar. Because data affected by autocorrelation are not independent (and potentially not identically distributed), conventional statistical analyses of such data may provide spurious results (e.g., inflated Type I error rates). The effect is similar to that caused by overestimating degrees of freedom, although the actual mechanism of the effect arises from incorrect partitioning of variances and covariances (Pagel 1993; see also Garland, Midford, and Ives 1999). Phylogenetically correlated data also lead to problems in parameter estimation and prediction (e.g., Martins and Hansen 1997; Garland and Ives 2000).

Several solutions to this problem have been proposed for cross-species comparisons (see recent reviews in Martins and Hansen 1996, 1997; Garland, Midford, and Ives 1999; Garland and Ives 2000). One of the most popular approaches is termed phylogenetically independent contrasts (Felsenstein 1985), which attempts to correct the data to remove the effect of phylogeny; in other words, transform quantitative data to create new values that are, in principle, independent and identically distributed. These values can then be used in conventional statistical procedures (see Garland, Harvey, and Ives 1992; Garland and Adolph 1994). The second approach, originally introduced by Martins and Garland (1991), is totally different. Instead of modifying the data to remove phylogenetic autocorrelation, the testing procedure is corrected to include the phylogenetic relationships in the statistical computations. This problem can be easily addressed with Monte Carlo simulations (Garland, Dickerman, Janis, and Jones 1993;

Reynolds and Lee 1996) in which the values of a trait are evolved along a phylogeny using a predetermined model of character evolution (e.g., Brownian motion). By repeating the process a large number of times for some fixed parameter computed from the actual data, or specified under the null hypothesis, one can easily obtain a null distribution of the reference statistic with which the actual one can be compared.

We propose another option for correcting the testing procedure. This method can be used with qualitative or quantitative traits. It is based on a permutation (randomization) procedure that accounts for the phylogeny. We will show how one can compute a permutation test for autocorrelated data, when the independence assumption is not satisfied. We will discuss the rationale and describe how the randomization procedure based on *phylogenetic permutations* (hereafter referred to as PP) actually works. A general algorithm for computing permutation tests under this model will be presented and compared to the more common *equally likely* (EL) model. The method will finally be applied to a published cross-species comparison to illustrate how it compares with a conventional statistical analysis.

2. Randomizations under the EL Model

Since R. A. Fisher (1935, Section 21), permutation tests have been used extensively in ecological and evolutionary studies (e.g., Sokal 1979; Dietz 1983; Douglas and Endler 1982; Dow and Cheverud 1985; Dow, Cheverud, and Friedlander 1987; Crowley 1992, pp. 407-409). The procedure is straightforward (but inadequate in the case of comparative data). Consider the case of testing for a correlation between two traits across a series of species. (a) First, one computes a statistic of interest based on actual observed values. (b) Then, the values of one trait are permuted while keeping the other fixed (by so doing, the correlation structure between traits is broken). (c) The same statistic as in Step 1 is computed for the permuted data. (d) This process (Steps b and c) is repeated a large number of times to generate a distribution of the test statistic. (e) The probability of the data under the null hypothesis is obtained by counting the number of test statistics (e.g., correlation values) for permuted sets of data which are larger than or equal to the original value of the statistic for the unpermuted data.

In theory, all possible $n!$ permutations among n species could be considered, but in practice, these tests are usually based on a large subset of permutations and are termed "sampled" randomization tests. Usually, the sampling universe is based on the EL model, which considers all permutations as equally likely, every possible permutation order of the trait values having the same probability ($1/n!$). However, Oden and Sokal (1992) have shown that in specific cases of permutation tests for distance matrices

(Mantel 1967; Hubert and Schultz 1976), the EL assumption can be inadequate. With simulations, Oden and Sokal observed that some of the methods designed to test for the partial correlation among three matrices (Dow and Cheverud 1985; Hubert 1985; Manly 1986; Smouse, Long, and Sokal 1986) produced inflated Type I error rates (see also Oden 1992). In particular, this problem can occur when the observations are autocorrelated.

To account for this problem, a modified permutation procedure is required. We want values for closely related species to be more likely permuted with each other, rather than being permuted equiprobably. One partial solution is to randomize the values within certain blocks, corresponding to one or more taxonomic levels (e.g., orders, families, genera), while preventing permutations across blocks (see Biondini, Mielke, and Berry 1988; Harvey and Pagel 1991, pp. 152-153). This approach is not general, however, because it does not account for the hierarchical phylogenetic relationships within blocks (taxa), nor for relationships of the blocks with each other. To be general, we require a completely specified phylogenetic tree, which in many cases will be entirely bifurcating and may have arbitrary branch lengths. That tree could then be used to compute a randomization test with phylogenetic permutations (PP).

3. Randomizations under the PP Model

Any phylogenetic tree can be uniquely depicted in the form of a path-length matrix containing the pairwise distances among all species (Hartigan 1967; Buneman 1971; Lapointe and Legendre 1991). The basic idea of the method is to convert the path-length phylogenetic distance matrix to a permutation probability (or transition) matrix **P**. This matrix gives for each species i (row) the probability p_{ij} of moving its trait value to a different species j (column), including itself (p_{ii}). Under the EL model, all probabilities for any given row would equal $1/n$, including self-permutations (i.e., transition probabilities to move the value from species i to species i). In the case of phylogenetic permutations (PP), this condition will not hold, however, and the transition probability matrix **P** will meet a different set of properties:

1. All transition probabilities p_{ij} should be positive.
2. All transition probabilities on any row i must sum to one.
3. The transition probabilities p_{ij} should be inversely and monotonically related to the path-length distances d_{ij} for any given species (row) i.
4. The self-permutation p_{ii} should always represent the largest transition probability for any given species (row) i.

5. The self-permutation for species that lack close relatives should be larger than for species with close relatives in the phylogenetic tree.

Although the first four properties seem obvious, the last one deserves an explanation. That property states that the probability of self-permutations will not be identical for all species and that the more distant and isolated a species is from the other species, the larger its self-permutation probability (i.e., the probability to remain in the same position) will be. This property differentiates species on long terminal branches from species on short terminal branches, as a Brownian motion model does in the case of Monte Carlo simulations (Martins and Garland 1991).

Obviously, the phylogenetic permutations depend on the path-length distances in **D**, and the transition probabilities in **P** can be obtained in several ways, so long as all properties are satisfied. We will now propose a simple procedure for computing a class of transition probability matrices from path-length distances (see Appendix). The method first proceeds by scaling all distances (d_{ij}) between zero and one, by dividing each path-length by the largest distance (d_{max}) in the matrix. The scaled distances (d^*_{ij}) are then converted to similarities (s_{ij}), where $s_{ij} = k - d^*_{ij}$, and k is any positive number larger than or equal to one. The similarities are finally transformed into transition probabilities (p_{ij}) by dividing each similarity value by its corresponding row total such that all marginals sum to one (see Property 2). These so-called "initial transition probabilities" that satisfy all properties are used for the PP algorithm. Depending on the value of parameter k, different matrices **P** will obtained, however. The larger k is, the more similar the probabilities in **P** will be, such that when $k = \infty$, **P** becomes the EL matrix. Therefore, the EL model is a special case of the PP model with all transition probabilities equal, and parameter k defines how far the transition matrix **P** is from the EL model.

4. The PP Algorithm

Unlike a standard permutation test for which the trait values for different species would be exchanged entirely at random and equiprobably (see Edgington 1995, pp. 3-5; Manly 1997, pp. 3-13), a transition probability matrix is used to guide the phylogenetic permutations. In particular, the procedure takes as input the trait value of one species at a time and returns a new address for it, until trait values for all species have been permuted. The major difference between the EL and the PP model is that values for different species are permuted with different probabilities. The detailed algorithm is as follows:

Step 1. Pick one species at random from the set of n species to be permuted.
Step 2. Assign a new position to this species according to the initial transition probability matrix.
Step 3. Pick a new object not already assigned to a position.
Step 4. Assign a new position to this species (one not already occupied by another species) according to the transition probabilities; to do so, the transition matrix must be reduced by deleting one row (the previously selected species) and one column (the new position of that species), and new probabilities must be computed so that they all sum to one.
Step 5. Repeat Steps 3 and 4 until only one species remains, and assign the trait values for that last species to the only position not yet occupied.

Using the EL model, the probability of any permutation (i.e., going from ABC to BAC) would be computed as $1/n!$ — but with the PP model, the probability of this same permutation is given by the product of the (successively computed) probabilities to move from A to B, from B to A, and from C to C. It is theoretically possible to compute the probabilities of all permutation orders with such a procedure; they should sum to one. However, because there exist several ways to obtain a particular order (e.g., by first moving from C to C, then from A to B, and from B to A) and because these may not have equal probabilities, it rapidly becomes intractable to compute realized or "final transition probabilities" as the number of species increases. These values are obviously obtained asymptotically by the PP algorithm when a large number of iterations are performed (see Appendix), and statistical tests can be computed to verify that they do not indeed differ statistically from the final transition probabilities.

5. Application of the PP Model

Huey and Bennett (1987) studied the thermal biology of 12 species of Australian scincid lizards, seeking to determine whether various aspects of thermal biology had evolved in a positively correlated fashion. Those authors measured four traits, but we will consider only two here: preferred body temperatures in a thermal gradient (a behavioral trait) and the optimal temperature for sprint running on a photocell-timed racetrack (a physiological trait). The authors hypothesized that lizards preferring high temperatures should also run fastest at high temperatures. The ordinary

Pearson product-moment correlation coefficient between preferred and optimal body temperatures (see Figure 1) is 0.5845; the conventional, one-tailed critical value for $\alpha = 0.05$ is 0.497. Based on these values, the correlation would be considered significant at $p = 0.0230$.

To establish a conventional null distribution of the Pearson correlation coefficient, for comparison with the permutation distributions, we used the PDSIMUL program (Garland et al. 1993) on a star phylogeny (no hierarchical structure). Using PDSIMUL on a star phylogeny is equivalent to drawing, for each species, a single pseudorandom number from a bivariate normal distribution of constant variance, and yields a distribution of correlation coefficients that is not significantly different from the conventional tables. For each of 5000 simulated data sets, we used the PDTIPS program (both programs are available from http://www.wisc.edu/zoology/faculty/fac/ Garland/PDAP.html) to compute the Pearson product-moment correlation coefficient.

The new PDRANDOM program (also available from T.G.) was used to do the permutation tests. First, we performed equally likely permutations and computed correlations with PDTIPS. We then compared their distributions with those for data simulated along a star phylogeny to detect significant differences, if any. To perform phylogenetic permutations, we used the phylogeny shown in Figure 1 to compute initial transition probabilities. For each set of 5000 permutations, we allowed the parameter k to vary. As is the convention in permutation tests, the original, unpermuted data were always included as one of the 5000 sets.

Table 1 shows the final transition probabilities of phylogenetic permutations under an EL model (with $k = \infty$) and a PP model with k varying from 2.0 to 1.0. Under an EL model, the value for each species is permuted to each position 8.3% (1/12) of the time, on average. Under all PP models shown, permutations occur unequally, depending on the value of parameter k. All final transition probabilities satisfy the five properties of the PP method as listed above, however.

Table 2 summarizes information on the distributions of the correlation coefficient computed from the simulated or permuted data sets. We used a t-test to compare means of the distributions and a Kolmogorov-Smirnov (KS) test to compare the distributions more generally. As expected, Monte Carlo simulations along a star phylogeny produced distributions of correlation coefficients that were virtually identical to the conventional ones. We next compared distributions of correlation coefficients under EL and PP permutations. With $k = 2.0$, the PP distributions of correlation values did not differ significantly from that generated under the EL model. With $k = 1.1$, however, the PP distributions differed significantly. Correspondingly, the significance levels were reduced (p-values were increased). With $k = 1.01$ or

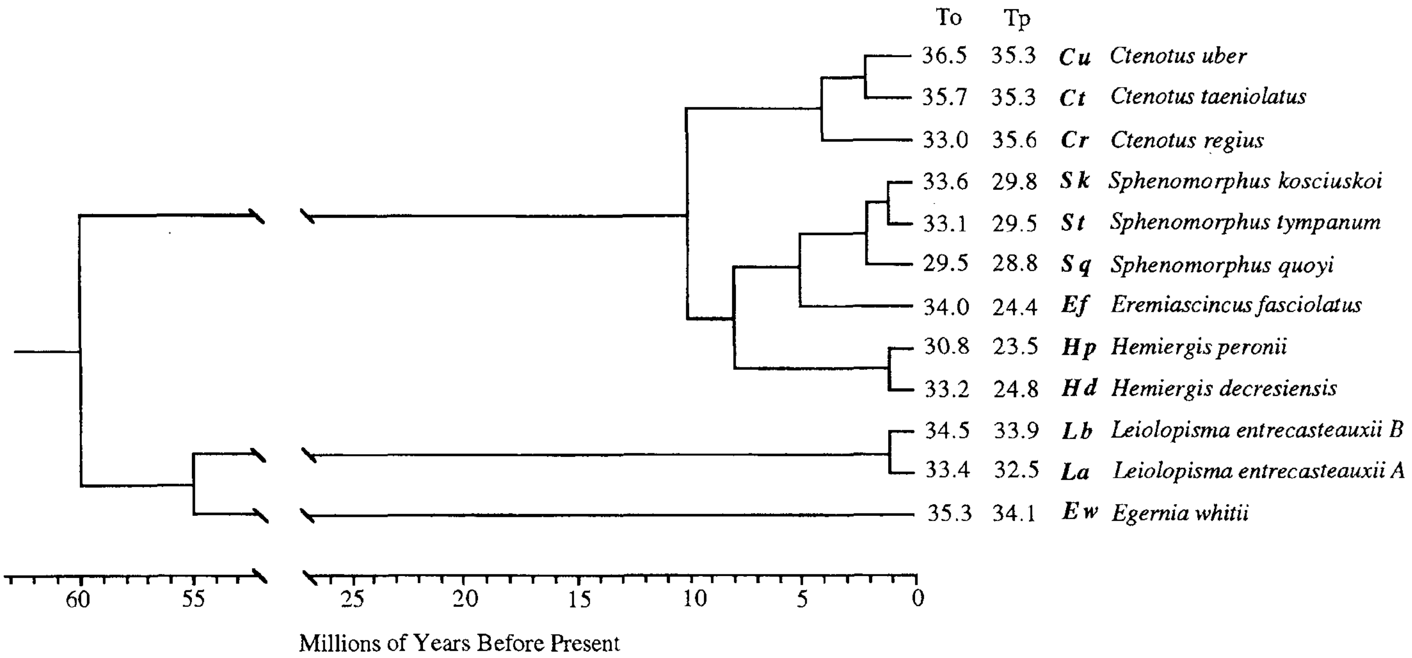


Figure 1. Hypothesized phylogenetic relationships and thermal data (^{o}C) for 12 species of Australian scincid lizards (from Fig. 1 and Table 1 of Garland, Huey, and Bennett 1991). T_o is optimal body temperature for sprint running, T_p is preferred body temperature.

Table 1. Examples of transition probabilities computed from the phylogenetic tree in Fig 1[a].

Equally Likely (EL) Permutations

	ew	la	lb	hd	hp	ef	sq	st	sk	cr	ct	cu
ew	8.9	8.3	8.1	7.2	8.1	8.7	8.5	9.1	8.1	8.7	7.9	8.5
la	8.1	7.5	8.6	8.8	8.2	8.6	8.3	8.0	8.5	8.6	8.9	7.8
lb	8.1	8.7	7.9	8.6	8.7	8.1	8.0	8.2	8.6	8.0	8.3	8.8
hd	8.1	8.5	8.4	8.5	8.7	7.8	7.8	8.8	8.5	8.3	8.2	8.4
hp	8.2	8.4	8.1	8.9	7.9	9.0	8.3	8.4	8.3	8.2	8.1	8.2
ef	8.1	7.8	8.0	8.8	8.8	8.1	8.9	7.9	8.1	8.9	8.2	8.3
sq	7.9	8.2	8.1	7.7	9.1	8.9	9.3	8.3	8.4	7.2	8.2	8.7
st	8.9	8.4	8.6	8.0	8.1	8.2	7.9	8.4	8.7	8.6	8.4	7.9
sk	8.4	9.6	7.9	8.1	8.6	7.7	8.7	8.2	7.5	9.2	8.3	7.7
cr	8.1	8.5	8.9	8.3	8.2	8.1	8.2	8.3	8.5	8.4	8.1	8.5
ct	8.4	8.0	8.9	8.3	7.8	8.2	8.0	8.3	8.1	8.3	8.9	8.7
cu	8.7	8.1	8.5	8.8	7.8	8.5	8.1	8.2	8.6	7.6	8.6	8.4

Phylogenetic Permutations (PP) with $k = 2.0$

	ew	la	lb	hd	hp	ef	sq	st	sk	cr	ct	cu
ew	18.6	10.1	9.8	6.9	6.6	6.8	6.6	7.0	7.1	7.0	7.0	6.6
la	9.8	15.4	16.1	6.5	6.4	6.2	6.4	6.6	7.4	6.4	6.5	6.5
lb	9.9	15.5	14.4	6.9	7.1	6.8	6.9	6.8	6.0	7.1	6.5	6.1
hd	7.6	6.4	6.3	9.2	9.5	9.8	8.4	8.3	8.0	8.7	8.8	9.0
hp	6.1	7.7	6.5	10.3	9.0	8.5	9.2	8.6	8.4	8.5	8.6	8.5
ef	6.6	6.4	6.6	9.2	9.0	9.3	9.6	8.7	8.3	8.8	9.3	8.4
sq	6.8	6.0	6.8	8.6	8.6	9.1	8.9	8.9	10.1	8.8	8.2	9.2
st	6.8	6.3	6.9	8.4	9.4	8.9	9.3	9.5	9.0	8.7	8.8	7.8
sk	6.8	6.4	6.9	8.5	8.7	8.7	9.0	9.3	9.4	8.5	8.5	9.1
cr	7.3	6.6	6.6	8.7	7.9	8.7	8.7	8.3	9.2	9.0	9.6	9.5
ct	7.3	6.7	6.3	8.6	9.0	8.5	8.3	9.0	8.6	9.0	9.4	9.4
cu	6.5	6.5	6.8	8.2	8.7	8.7	8.8	9.0	8.6	9.5	8.8	9.9

(Table 1 continues on next page)

k = 1.0 (in which case permutations across the root of the phylogeny are precluded; see bottom of Table 1), distributions of correlations under PP models differed even more from the EL distributions. Significance levels were reduced further, and in both cases the Pearson correlation would not be declared significant at $\alpha = 0.05$.

In summary, as shown in Table 1, the PP algorithm implemented in the PDRANDOM program succeeds in producing randomizations that reflect phylogenetic structure and satisfy the above-listed desirable properties of the method. Depending on the value of the parameter k, distri-

Table 1 (*continued*)

Phylogenetic Permutations (PP) with $k = 1.1$

	ew	la	lb	hd	hp	ef	sq	st	sk	cr	ct	cu
ew	53.9	8.4	8.2	3.1	3.8	3.4	2.9	3.5	3.3	3.1	3.1	3.3
la	8.3	33.4	34.5	2.5	2.4	2.6	2.7	2.6	2.4	2.7	2.8	3.1
lb	7.9	34.3	34.1	2.7	2.8	2.8	2.9	2.4	2.6	2.5	2.4	2.6
hd	3.3	2.4	2.5	10.7	10.5	10.1	10.3	9.7	10.6	9.7	9.9	10.4
hp	3.7	2.5	2.7	11.7	10.5	9.8	9.9	10.3	9.6	9.0	9.9	10.3
ef	3.1	2.5	2.7	10.0	10.3	11.2	11.0	10.5	9.2	9.6	10.6	9.1
sq	3.0	2.6	2.3	10.1	9.7	10.4	10.6	11.0	10.7	9.3	10.6	9.5
st	3.0	2.8	2.5	10.2	10.3	10.4	10.1	10.9	10.3	10.1	9.8	9.5
sk	3.6	3.2	2.5	9.8	9.4	10.5	10.7	11.1	11.4	9.1	9.3	9.4
cr	3.7	2.6	2.9	9.4	10.1	9.4	10.5	8.9	9.7	12.0	10.5	10.3
ct	3.2	2.6	2.4	10.1	10.4	9.5	9.1	9.6	9.5	11.2	10.9	11.4
cu	3.3	2.6	2.8	9.8	9.7	9.9	9.2	9.2	10.6	11.5	10.3	11.0

Phylogenetic Permutations (PP) with $k = 1.01$

	ew	la	lb	hd	hp	ef	sq	st	sk	cr	ct	cu
ew	81.8	6.4	6.6	0.8	0.6	0.6	0.3	0.6	0.5	0.7	0.6	0.5
la	7.1	43.0	45.3	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.4	0.5
lb	5.5	46.7	43.9	0.5	0.4	0.4	0.5	0.4	0.4	0.4	0.6	0.4
hd	0.7	0.5	0.5	11.8	13.0	10.1	10.6	10.3	10.7	10.9	10.9	10.2
hp	0.6	0.5	0.5	12.2	12.6	10.3	11.3	10.3	11.1	9.6	10.5	10.5
ef	0.6	0.6	0.4	10.4	10.2	12.4	11.2	11.1	10.7	11.4	10.4	10.6
sq	0.7	0.4	0.4	10.3	10.8	11.2	11.9	12.1	11.1	11.0	9.9	10.2
st	0.9	0.5	0.4	10.8	10.1	11.8	10.9	11.4	12.2	10.2	11.1	9.9
sk	0.5	0.3	0.6	9.5	10.0	12.7	11.1	11.8	11.8	10.5	10.4	10.9
cr	0.4	0.4	0.7	10.8	10.5	9.8	10.7	11.1	10.6	12.1	11.4	11.4
ct	0.5	0.4	0.4	11.1	10.8	9.6	10.7	10.4	10.4	11.5	11.7	12.4
cu	0.6	0.5	0.4	11.4	10.5	10.6	10.4	10.1	9.9	11.1	12.1	12.5

(Table 1 continues on next page)

butions of the Pearson correlation coefficient computed from the permuted data sets differ significantly from those computed for equally likely (EL) permutations (Table 2). For the real data set shown in Figure 1, the choice of k has an important effect on the results: depending on the value of k, the correlation would be judged not statistically significant at $\alpha = 0.05$, whereas it would be judged significant under an EL permutation model. This pattern is consistent with the findings of Garland, Huey, and Bennett (1991); phylogenetic analyses using either independent contrasts or Monte Carlo computer simulations both failed to reject the null hypothesis.

Table 1 (*continued*)

Phylogenetic Permutations (PP) with $k = 1.0$

	ew	la	lb	hd	hp	ef	sq	st	sk	cr	ct	cu
ew	88.3	5.5	6.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
la	5.7	48.3	46.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
lb	6.1	46.2	47.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
hd	0.0	0.0	0.0	12.4	11.5	11.2	10.9	10.3	11.3	12.0	10.7	9.8
hp	0.0	0.0	0.0	12.6	13.1	10.6	11.0	11.2	11.0	9.8	10.1	10.5
ef	0.0	0.0	0.0	10.4	11.1	12.2	11.6	11.3	11.5	10.4	10.6	10.8
sq	0.0	0.0	0.0	10.7	10.4	11.6	11.9	12.8	11.7	11.2	9.4	10.3
st	0.0	0.0	0.0	11.8	10.7	11.4	11.8	12.2	11.1	10.6	10.8	9.5
sk	0.0	0.0	0.0	10.2	11.2	11.7	11.5	11.5	11.8	10.3	10.7	11.1
cr	0.0	0.0	0.0	10.8	10.8	10.4	10.6	10.4	10.6	12.3	11.8	12.3
ct	0.0	0.0	0.0	10.2	10.6	10.7	10.2	10.0	10.1	11.9	13.4	12.9
cu	0.0	0.0	0.0	10.9	10.5	10.2	10.5	10.3	10.8	11.6	12.3	12.7

[a] Rows are source tips, columns are destinations. 5000 permutations were performed: values are final transition probabilities for trait 1. Self-permutations are underlined.

6. Discussion

In this paper, we have proposed a new phylogenetic permutation method to analyze cross-species data while accounting for phylogenetic nonindependence. The phylogenetic permutation (PP) procedure is a generalization of the equally likely (EL) model used in standard randomization tests (Oden and Sokal 1992). The main difference between the PP and EL models is the permutation procedure used to compute the probability of the data under the null hypothesis. Thus, our method can correct for the inflated Type I errors that occur when phylogenetically nonindependent observations are analyzed with conventional statistical methods. Note, however, that the PP procedure does not correct for any bias in the estimation of statistical parameters (e.g., the slope of a regression). Other approaches, such as independent contrasts (Felsenstein 1985) or generalized least squares (Grafen 1989; Martins and Hansen 1997; Garland and Ives 2000), can be used to improve parameter estimation *per se*, and the PP procedure (like Monte Carlo simulations, Martins and Garland 1991; Garland et al. 1993; Reynolds and Lee 1996) could be used in combination with these methods. Future studies will be required to compare the statistical properties of the PP procedure with those of Monte Carlo simulations and other comparative methods for cross-species data.

Table 2. Summary statistics for distributions of correlation coefficients[a].

	k	Mean r	Percentiles 2.5	5	95	97.5	p[b]
Conventional Test :					0.497	0.576	0.0230
Monte Carlo Simulations[c]:	n/a	-0.001	-0.565	-0.492	0.501	0.574	0.0228
EL Permutations[c]:	n/a	0.001	-0.556	-0.480	0.514	0.584	0.0244
Phylogenetic Permutations:	2.00[d]	0.009	-0.547	-0.477	0.509	0.596	0.0284
	1.10[e]	0.047	-0.535	-0.463	0.559	0.630	0.0384
	1.01	0.094	-0.485	-0.414	0.608	0.663	0.0636
	1.00	0.108	-0.466	-0.406	0.620	0.682	0.0766

[a] PDSIMUL (Garland et al. 1993) was used to perform Monte Carlo simulations; PDRANDOM (this paper) was used to perform permutations (see Table 1 for final transition probabilities). In both cases, PDTIPS was used to analyze the data. 5000 permutations were performed. 95th percentiles are underlined as these values would be used to test (one-tailed) against the null hypotheses for $\alpha = 0.05$. The correlation for the real data set (shown in Fig. 1) was 0.5845.

[b] p = number of randomized (or simulated) correlation coefficients that are greater than or equal to the real one, divided by 5000. Note that PDRANDOM includes the real data set as the first one in its output (*.RND) file.

[c] Distributions do not differ significantly: $t = -0.42$, d.f. = 9,998, 2-tailed $p = 0.673$ and K-S Z = 0.610, 2-tailed $p = 0.851$.

[d] Distributions do not differ significantly from distribution for EL Permutations: $t = -1.27$, 2-tailed $p = 0.203$ and K-S Z = 1.130, 2-tailed $p = 0.155$.

[e] Distributions differ significantly from distribution for EL Permutations: (Levene's test for unequal variances F = 5.50, $p = 0.019$; $t = -7.51$, unequal variance d.f. = 9,989.52, 2-tailed p = < 0.0005; K-S Z = 3.630, 2-tailed $p < 0.0005$).

Phylogenetic permutations represent models that can be applied to a wide range of problems. Indeed, our approach is more general than most other tests designed for correcting Type I error in comparative studies. Unlike the independent contrasts method, it is not the trait values themselves that are changed with our procedure, but the positions of the trait values at the "leaves" (tips) of a phylogenetic tree. Thus, phylogenetic permutations can be performed with any kind of data, qualitative as well as quantitative. Also, because prior information about various biological parameters, such as limits to trait evolution (Garland et al. 1993; Reynolds and Lee 1996), is not required, phylogenetic permutations are easier to implement than phylogenetic simulations. In our procedure, the only information required, apart from the trait values, is an estimate of the phylogeny, from which a transition matrix **P** can be computed. We proposed a method to obtain **P** from a matrix **D** of path-length distances, but any matrix containing initial transition probabilities could be used to permute the data. Similarly, the parameter k can be varied to assess the sensitivity of test data to different levels of phylogenetic permutations, including the EL model ($k = \infty$). As we have shown in the example (Table 1), the outcome of the test can be greatly affected by this value k. An interesting area for future work would be to determine if and how different values of k relate to different evolutionary models (see Martins and Hansen 1997). An equally important area would be the comparison of the statistical properties of alternative methods for computing matrix **P** (e.g., taking the inverse of path-length distances as initial transition probabilities).

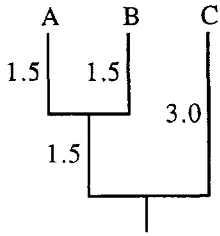
Limited or constrained permutations can also be performed with the PP procedure to assess any conditional hypothesis. For instance, one could prevent the values for one group of species from changing positions with another group by setting the relevant values in **P** to zero; this strategy would create two blocks of species, allowing permutations to be performed within each group, but not among them (e.g., see Table 2, for $k = 1.0$). Similarly, one can force a species to remain in its original position to control the effect of that particular species (or a few species) on the outcome of the test; practically, this procedure is done by setting the self-permutation probability to one for the selected species, and all other probabilities in that row to zero. The flexibility of our procedure thus allows the user to address the statistical implications of specific phylogenetic hypotheses. For example, a stepwise approach could be performed by first computing a global test allowing all permutations, followed by *a posteriori* testing that would prevent the permutations among some predetermined groups. So long as a corresponding transition matrix **P** can be obtained, permutations may be performed in accordance with any given model. A final generalization of our approach is the possibility of using different transition matrices for different traits. The EL model is a special case, which always uses the same matrix for each trait, or only permutes a single trait (which is functionally equivalent).

When several competing phylogenies are available for a given set of species (see Kirsch and Lapointe 1997), phylogenetic permutations can be used to compare the outcome of the tests computed with different transition matrices. However, when the phylogeny is incomplete or unknown, this randomization of the trait positions on a tree can be combined with some other randomization methods for trees themselves. For example, one could generate a series of trees with random topologies and randomly assigned branch (edge) lengths. Then, either phylogenetic permutations or simulations (Losos 1994; Abouheif 1998) could be applied to each phylogeny to assess the generality of the statistical outcome (see also Blomberg 2000). When the topology of the phylogeny is known, but actual branch lengths are unknown, a transition probability matrix satisfying topological relationships only can also be computed. To do so, one may use either path-length distances with arbitrary branch lengths, such as all branch lengths set to be equal (for one of many alternatives, see Grafen 1989), or simply count the number of nodes separating two species (for other possibilities, see Podani and Dickinson 1984; Podani 2000). When using topological distances, the transition matrices for each trait in an analysis are necessarily the same. When using path-length distances, however, they need not be. Thus, we could use different sets of branch lengths to derive the transition matrices for different traits as is often done with independent contrasts (e.g., see Bonine and Garland 1999).

Although originally motivated by the problem of phylogenetic correlation, our procedures can also be applied to other types of autocorrelation. For spatially autocorrelated data (Clifford, Richardson, and Hémon 1989; Legendre 1993), a permutation model could be implemented according to the geographic distances among the various localities at which species have been sampled. The rationale in this particular case would be that contiguous sampling sites are more likely to share the same species than distant sites (positive autocorrelation), or the opposite (negative autocorrelation, perhaps caused by competition). The same procedure can be used with time series in which a permutation matrix **P** representing an autoregressive model would represent the different times at which species were collected. In that case, successive sampling dates will be more likely to be permuted than distant ones. Also, it is possible to include more than one **P** matrix in the analyses to permute species according to competing models representing different sources of autocorrelation (Dutilleul and Potvin 1995).

Appendix: Computation of Final Transition Probabilities

Consider the simple case of the smallest possible tree with three leaves.



$$\mathbf{D} = \begin{array}{c|ccc} & A & B & C \\ \hline A & 0.0 & 3.0 & 6.0 \\ B & 3.0 & 0.0 & 6.0 \\ C & 6.0 & 6.0 & 0.0 \end{array}$$

The initial transition probability matrix **P** is obtained from the corresponding path-length distance matrix **D**, with $k = 2.0$ (see Section 3).

	A	B	C
A	0.444	0.333	0.222
B	0.333	0.444	0.222
C	0.250	0.250	0.500

To compute the realized transition probabilities for each pair of species, we first need to compute the probability of each order (ABC, ACB, BAC, BCA, CAB, and CBA) and thus require the initial transition matrix and the reduced transition matrices computed after each permutation. For simplicity, first consider the probability of obtaining order ABC (i.e., A moves to A, B to B, and C to C). We get from **P** that the probability of moving from A to A (i.e., p_{AA}) is 4/9 (0.444). After reduction, we obtain that p_{BB} is 2/3 (O.667), and finally p_{CC} is obviously 1.000. Thus, the probability of obtaining the order ABC is 4/9 x 2/3 x 1 = 8/27 (0.296).

As another example, consider the order ACB. We need the values p_{AA}, p_{BC}, and p_{CB} to do the computations. p_{AA} is again 4/9 (0.444). The next transition probability (p_{BC}) is obtained from the reduced matrix and equals 1/3 (0.333). p_{CB} is equal to 1.000. Thus, the probability of the order ACB is given by: 4/9 x 1/3 x 1 = 4/27 (0.148). However, there are six distinguishable ways of getting this order, each depending on the order in which the species (rows) are selected. In the example presented above, the selection order of the species was ABC and the permutation result was ACB. Now, what if we had picked species C first, then B, and finally A? That is, we want to compute the probability or getting the permutation order ACB by considering the selection order CBA. In this case, we find that ACB is obtained with probability 1/4 x 2/5 x 1 = 1/10 (0.100). This value differs from the 4/27 (0.148) we obtained with the selection order ABC. That demonstration implies that all possible selection orders must be considered to compute the probability of a given permutation order.

Here are the frequencies of each possible permutation order (rows) for a given selection order (columns). The sum of the columns represents the cumulative frequency for a given selection order (all equal to one), whereas the sum of the rows represent the cumulative frequency for each permutation order. The probabilities of permutation orders (relative frequencies) are also provided for each row:

	ABC	ACB	BAC	BCA	CAB	CBA	TOTAL	PROB
ABC	8/27	8/27	8/27	8/27	2/7	2/7	1.7566	0.2928
ACB	4/27	4/27	8/63	1/9	1/6	1/10	0.8011	0.1335
BAC	1/5	2/9	1/5	2/9	3/14	3/14	1.2730	0.2122
BCA	2/21	1/9	2/15	1/9	1/12	3/20	0.6841	0.1140
CAB	2/15	1/9	2/21	1/9	3/20	1/12	0.6841	0.1140
CBA	8/63	1/9	4/27	4/27	1/10	1/6	0.8011	0.1335
TOTAL	1.00	1.00	1.00	1.00	1.00	1.00		

The final transition probability matrix from species i to j is obtained by adding the relevant probabilities in the permutation order matrix. For example, p_{AC} is computed as 0.1140 + 0.1335. The complete matrix is thus:

	A	B	C
A	0.4263	0.3262	0.2475
B	0.3262	0.4263	0.2475
C	0.2475	0.2475	0.5049

Given this consideration, it becomes cumbersome to compute final transition probabilities, because we first need to compute permutation order probabilities, which are themselves affected by selection orders. It can be done for such small matrices, as we have shown here. However, this procedure is practically impossible for larger matrices. Nevertheless, the algorithm we are using will asymptotically recover the correct probabilities. Indeed, with 5000 permutations, we have obtained the following empirical results that approach the expected probability values. These observed permutations satisfy all five properties of the PP algorithm (see Section 3).

	A	B	C
A	0.4360	0.3260	0.2380
B	0.3240	0.4280	0.2480
C	0.2400	0.2460	0.5140

References

ABOUHEIF, E. (1998), "Random Trees and the Comparative Method: A Cautionary Tale," *Evolution, 52,* 1197-1204.

BIONDINI, M. E., MIELKE, P. W., and BERRY, K. J. (1988), "Data-Dependent Permutation Techniques for the Analysis of Ecological Data," *Vegetatio*, *75*, 161-168.

BLOMBERG, S. P. (2000), "Fels-Rand: An Xlisp-Stat Program for the Comparative Analysis of Data Under Phylogenetic Uncertainty," *Bioinformatics*, *16,* 1010-1013.

BONINE, K. E., and GARLAND, T., JR. (1999), "Sprint Performance of Phrynosomatid Lizards, Measured on a High-Speed Treadmill, Correlates with Hindlimb Length," *Journal of Zoology, London, 248,* 255-265.

BROOKS, D. R., and MCLENNAN, D. A. (1991), *Phylogeny, Ecology, and Behavior. A Research Program in Comparative Biology,* Chicago: University of Chicago Press.

BUNEMAN, P. (1971), "The Recovery of Trees From Measures of Dissimilarity," in *Mathematics in Archeological and Historical Sciences,* Eds. F. R. Hudson, D.G. Kendall, and P. Tautu, Edinburgh: Edinburgh University Press, 387-395.

CLIFFORD, P., RICHARDSON, S., and HEMON, D. (1989), "Assessing the Significance of the Correlation Between two Spatial Processes," *Biometrics, 45,* 123-134.

CROWLEY, P. H. (1992), "Resampling Methods for Computation-Intensive Data Analysis in Ecology and Evolution," *Annual Review of Ecology and Systematics, 23,* 405-447.

DIETZ, E. J. (1983), "Permutation Tests for Association Between Distance Matrices," *Systematic Zoology, 32,* 21-26.

DOUGLAS, M. E., and ENDLER, J. A. (1982), "Quantitative Matrix Comparisons in Ecological and Evolutionary Investigations," *Journal of Theoretical Biology, 99,* 777-795.

DOW, M. M., and CHEVERUD, J. M. (1985), "Comparison of Distance Matrices in Studies of Population Structure and Genetic Microdifferentiation: Quadratic Assignment," *American Journal of Physical Anthropology, 68,* 367-373.

DOW, M. M., CHEVERUD, J. M., and FRIEDLANDER, J. S. (1987), "Partial Correlation of Distance Matrices in Studies of Population Structure," *American Journal of Physical Anthropology, 72,* 343-352.

DUTILLEUL, P., and POTVIN, C. (1995), "Among-Environment Heteroscedasticity and Genetic Autocorrelation: Implications for the Study of Phenotypic Plasticity," *Genetics, 139,* 1815-1829.

EDGINGTON, E. S. (1995), *Randomization Tests* (3rd Ed.), New York: Marcel Dekker.

EGGLETON, P., and VANE-WRIGHT, R. I. (Eds.) (1994), *Phylogenetics and Ecology,* London: Academic Press.

FELSENSTEIN, J. (1985) "Phylogenies and the Comparative Method," *American Naturalist, 125,*1-15.

FISHER, R. A. (1935), *Design of Experiments,* Edinburgh: Oliver and Boyd.

GARLAND, T., JR., and ADOLPH, S. C. (1994), "Why not to do Two-Species Comparative Studies: Limitations on Inferring Adaptation," *Physiological Zoology, 67,* 797-828.

GARLAND, T., JR., DICKERMAN, A. W., JANIS, C. M., and JONES, J. A. (1993), "Phylogenetic Analysis of Covariance by Computer Simulation," *Systematic Biology, 42,* 265-292.

GARLAND, T., JR., HARVEY, P. H., and IVES, A. R. (1992), "Procedures for the Analysis of Comparative Data Using Phylogenetically Independent Contrasts," *Systematic Biology, 41,* 18-32.

GARLAND, T., JR., HUEY, R. B., and BENNETT, A. F. (1991), "Phylogeny and Coadaptation of Thermal Physiology in Lizards: A Reanalysis," *Evolution, 45,* 1969-1975.

GARLAND, T., JR., and IVES, A. R. (2000), "Using the Past to Predict the Present: Confidence Intervals for Regression Equations in Phylogenetic Comparative Methods," *American Naturalist, 155,* 346-364.

GARLAND, T., JR., MIDFORD, P. E., and IVES A. R. (1999), "An Introduction to Phylogenetically Based Statistical Methods, With a New Method for Confidence Intervals on Ancestral Values," *American Zoologist, 39,* 374-388.

GRAFEN, A. (1989), "The Phylogenetic Regression," *Philosophical Transactions of the Royal Society of London- Series B: Biological Sciences, 326,* 119-157.

HARTIGAN, J. A. (1967), "Representation of Similarity Matrices by Trees," *Journal of the American Statistical Association, 62,* 1140-1158.

HARVEY, P. H., and PAGEL, M. D. (1991), *The Comparative Method in Evolutionary Biology,* Oxford: Oxford University Press.

HUBERT, L. J. (1985), "Combinatorial Data Analysis: Association and Partial Association," *Psychometrika, 50,* 449-467.

HUBERT, L. J., and SCHULTZ, J. (1976), "Quadratic Assignment as a General Data Analysis Strategy," *British Journal of Mathematical and Statistical Psychology, 29,* 190-241.

HUEY, R. B., and BENNETT, A. F. (1987), "Phylogenetic Studies of Coadaptation: Preferred Temperatures Versus Optimal Performance Temperatures of Lizards," *Evolution, 41,* 1098-1115.

KIRSCH, J. A. W., and LAPOINTE, F.-J. (1997), "You Aren't (Always) What you Eat: Evolution of Nectar-Feeding Among Old World Fruitbats (Megachiroptera: Pteropodidae)," in *Molecular Evolution and Adaptive Radiation,* Eds. K. Sytsma and T. Givnish, New York: Cambridge University Press, 313-330.

LAPOINTE, F.-J. and LEGENDRE, P. (1991), "The Generation of Random Ultrametric Matrices Representing Dendrograms," *Journal of Classification, 178,* 177-200.

LEGENDRE, P. (1993), "Spatial Autocorrelation: Trouble or New Paradigm," *Ecology, 74,* 1659-1673.

LOSOS, J. B. (1994), "An Approach to the Analysis of Comparative Data When a Phylogeny is Unavailable or Incomplete," *Systematic Biology, 43,* 117-123.

MANLY, B. F. J. (1986), "Randomization and Regression Methods for Testing for Associations With Geographical, Environmental and Biological Distances Between Populations," *Researches in Population Ecology, 28,* 201-218.

MANLY, B. F. J. (1997), *Randomization, Bootstrap and Monte Carlo Methods in Biology* (2nd Ed.), London: Chapman and Hall.

MANTEL, N. (1967), "The Detection of Disease Clustering and a Generalized Regression Approach," *Cancer Research, 27,* 209-220.

MARTINS, E. P. (Ed.) (1996), *Phylogenies and the Comparative Method in Animal Behavior,* Oxford: Oxford University Press.

MARTINS, E. P., and GARLAND, T., JR. (1991), "Phylogenetic Analysis of the Correlated Evolution of Continuous Characters: A Simulation Study," *Evolution, 45,* 534-557.

MARTINS, E. P., and HANSEN, T. F. (1996), "The Statistical Analysis of Interspecific Data: A Review and Evaluation of Comparative Methods," in *Phylogenies and the Comparative Method in Animal Behavior,* Ed. E. P. Martins, Oxford: Oxford University Press, 22-75.

MARTINS, E. P., and HANSEN, T. F. (1997), "Phylogenies and the Comparative Method: A General Approach to Incorporating Phylogenetic Information Into the Analysis of Interspecific Data," *American Naturalist, 149,* 646-667. Erratum *153*, 448.

ODEN, N. L. (1992), "Spatial Autocorrelation Invalidates the Dow-Cheverud Test," *American Journal of Physical Anthropology, 89,* 257-264.

ODEN, N. L., and SOKAL, R. R. (1992), "An Investigation of Three-Matrix Permutation Tests," *Journal of Classification, 9,* 275-290.

PAGEL, M. (1993), "Seeking the Evolutionary Regression Coefficient: An Analysis of What Comparative Methods Measure," *Journal of Theoretical Biology, 164,* 191-205.

PODANI, J. (2000), "Simulation of Random Dendrograms and Comparison Tests: Some Comments," *Journal of Classification, 17,* 123-142.

PODANI, J., and DICKINSON, T. A. (1984), "Comparison of Dendrograms: A Multivariate Approach," *Canadian Journal of Botany, 62,* 2765-2778.

REYNOLDS, P. S., and LEE, R. M., III. (1996), "Phylogenetic Analysis of Avian Energetics: Passerines and Nonpasserines Do not Differ," *American Naturalist, 147,* 735-759.

SMOUSE, P. E., LONG, J. C., and SOKAL, R. R. (1986), "Multiple Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence," *Systematic Zoology, 35,* 627-632.

SOKAL, R. R. (1979), "Testing Statistical Significance of Geographical Variation Patterns," *Systematic Zoology, 28,* 227-232.