Chapter 2

Adaptation, Phylogenetic Inertia, and the Method of Controlled Comparisons

STEVEN HECHT ORZACK AND ELLIOTT SOBER

In the *Philosophical Investigations*, Wittgenstein (1968, 94) describes a man who is unconvinced by a claim he reads in a newspaper, so he buys another copy of the same newspaper to double-check. The point of the joke is that it isn’t just the quantity of evidence in support of a claim that matters, but the quantity of independent evidence (Sober 1989).

The issue of independent evidence is of central importance to hypothesis testing in evolutionary biology. Suppose you wanted to test the hypothesis that long fur is an adaptation to cold climate and short fur is an adaptation to warm climate. You look at 20 bear species; 10 live in a cold climate and have long fur, and 10 live in a warm climate and have short fur. Is there any reason to think that the data do not confirm the adaptive hypothesis? One worry is that the species in each group resemble each other merely because they inherited their fur length from a common ancestor of the group (and that the temperatures experienced by ancestors and descendants are similar). This influence of ancestor on descendant is often called phylogenetic inertia (e.g., see Harvey and Pagel 1991).

If phylogenetic inertia and natural selection are alternative explanations of the data, how are these competing hypotheses to be evaluated? The following quotation exemplifies a common idea about this problem:

Why do most land vertebrates have four legs? The seemingly obvious answer is that this arrangement is the optimal design. This response would ignore, however, the fact that the fish that were ancestral to terrestrial animals also have four limbs, or fins. Four limbs may be very suitable for locomotion on dry
land, but the real reason that terrestrial animals have this arrangement is
because their evolutionary predecessors possessed the same pattern. (Lewin
1980, 886)

Lewin’s claim is that if a trait can be explained by natural selection
and also by phylogenetic inertia, the latter explanation should be
regarded as a null hypothesis (see also Ridley 1983).

In this chapter, we clarify what phylogenetic inertia means and
discuss how adaptive and inertial hypotheses should be tested. We
argue that selection and inertia are not necessarily in conflict because
it is possible for neither, one, or both to significantly influence any
given trait, and that neither inertia nor selection deserves to be
regarded as a null hypothesis. We propose a new method for testing
these hypotheses, which we call the method of controlled compar-
isons. We also discuss the relationship between this method and the
Felsenstein’s method allows the investigator to control for the possi-
ble influence of phylogenetic inertia when the goal is to test an adap-
tive hypothesis. In contrast, our method allows one to test hypotheses
of inertia and hypotheses of selection simultaneously.

WHAT IS PHYLOGENETIC INERTIA?

What does it mean to say that land vertebrates have four limbs
because of phylogenetic inertia? This hypothesis goes beyond the
claim that there was no change in trait value. After all, adaptationists
can and often do contend that lack of change in a trait is due to sta-
bilizing selection and not inertia. By the same token, it is consistent
with the existence of inertia that a descendant’s trait value deviates
from the trait value of its ancestor. For these reasons, lack of change
in a trait is neither sufficient nor necessary evidence for the existence
of phylogenetic inertia (e.g., see Wake et al. 1983 and Hansen 1997).
This point highlights the importance of drawing a distinction
between pattern and process.

Phylogenetic inertia means that trait values of ancestors “influ-
ence” the trait values of descendants. What does this mean? The
following model of a discretely varying trait illustrates this influence.
The lessons we draw from it apply to continuously varying traits as
well. Consider a trait with two states: 0 and 1. A lineage in state 0 has
a probability \((1 - u)\) of being in state 0 after a given small interval of
time, and a lineage in state 1 has a probability \((1 - v)\) of being in state
1 after the same small interval of time; here, \(u\) and \(v\) are the net prob-
babilities of change stemming from all the evolutionary influences
on the lineage. These include natural selection, genetic drift, and
developmental or genetic constraints (as discussed by, for example,
can compute the probability, \(P_i(i \rightarrow j)\), that a lineage is in state \(j\)
at time \(t\) given that it was in state \(i\) at time 0. If \(u\) and \(v\) are small, the
standard theory of stochastic processes (e.g., Parzen 1962, 293–295)
indicates that

\[
P_i(0 \rightarrow 0) = \frac{v}{u+v} + \frac{u}{u+v} e^{-\lambda t}
\]

\[
P_i(0 \rightarrow 1) = \frac{v}{u+v} - \frac{u}{u+v} e^{-\lambda t}
\]

\[
P_i(1 \rightarrow 0) = \frac{u}{u+v} - \frac{v}{u+v} e^{-\lambda t}
\]

\[
P_i(1 \rightarrow 1) = \frac{v}{u+v} + \frac{u}{u+v} e^{-\lambda t}
\]

These equations for stasis and change of a trait show how the initial
state of a lineage as well as evolutionary forces such as natural selec-
tion exert an influence on its end state. The influence of the initial
state on the end state is phylogenetic inertia. If the lineage is young,
the effect of phylogenetic inertia is greater, whereas if the lineage
is old, it is smaller. In the infinite limit, \(P_i(0 \rightarrow 0) = P_i(1 \rightarrow 0) = \frac{v}{u+v}\)
and \(P_i(1 \rightarrow 1) = P_i(0 \rightarrow 1) = \frac{v}{u+v}\), meaning that the lineage has the
same probability of ending in a given state, regardless of its initial
state. Only at this limit is there no phylogenetic inertia. This result
underscores an important point: Most plausible models of trait evolu-
tion imply that traits in the real world always exhibit the influence
of their ancestors. What is at stake is the importance of this influence
and not whether phylogenetic inertia is present. Whether inertia is
found to significantly affect a particular trait depends on the kind of
statistical test used and the sample size involved. The following dis-
cussion ignores such concerns in that we refer to the presence and
absence of inertia in regard to a given trait. This talk is shorthand for
the contingent claim that ancestral trait values are statistically signif-
icient predictors of the trait values of extant species.
A specific example of inertia is provided by the Jukes and Cantor (1969) model of DNA sequence evolution. The trait is the identity of the nucleotide (A, T, C, or G) present at one site in a locus. For simplicity, we concentrate only on A and suppose that it has one fitness value and that T, C, and G have another. Let \( \alpha \) be the probability that A will change to any of the three other states during a small period of time, and let \( \beta \) be the probability that each of T, C, and G will change to A. The values of \( \alpha \) and \( \beta \) are jointly determined by the forces of natural selection and genetic drift. The equations for stasis and change for A, after a period of time \( t \), are

\[
P_t(A \to A) = \frac{\beta}{3\alpha + \beta} + \frac{\alpha}{3\alpha + \beta} e^{-(\alpha + \beta) t}
\]

and

\[
P_t(A \to \text{not-A}) = \frac{3\alpha}{3\alpha + \beta} - \frac{\alpha}{3\alpha + \beta} e^{-(\alpha + \beta) t}
\]

As before, these formulas specify an equilibrium value and a correction factor reflecting the influence of the initial condition, which shrinks with time. As this example illustrates, there is nothing mysterious about how phylogenetic inertia arises. Phylogenetic inertia is defined with respect to the history of a lineage, but it is a consequence of the generation-by-generation transmission of traits. In this example, one might test whether natural selection occurs and also whether inertia has a significant influence. We note in passing that, in principle, inertia could be tested alongside any hypothesis about the causes of trait values. One can think of instances in which it would naturally be tested along with the hypothesis of genetic drift.

A direct way to test for the influences of natural selection and phylogenetic inertia on a trait's evolution relies on knowing the net probabilities of change and the age of the species in question. Suppose, as before, that one examines a trait with states 0 and 1. Trait 0 is present in all of the extant species in a phylogeny, and it is known to be the ancestral condition. The values of \( u \) and \( v \) bear on the acceptance of the hypotheses under test in the way shown in Table 2.1. If \( u < v \), then selection favored the retention of state 0. If \( u \geq v \), the opposite is true. If \( t \) is "small," the state of the ancestor exerted a significant influence on the state of descendants, and so phylogenetic inertia was present. If \( t \) is "large," the ancestor did not exert such an influence. Suppose that \( u = 0.0001 \) and \( v = 0.001 \). This means that selection was one of the causes of the stasis we observe. Did inertia also contribute significantly? The problem is to decide whether the absolute time elapsed is small or large. One possible way to decide this issue is to determine whether the observed frequency of trait 0 (100%) is significantly greater than 91% \((=100 \times \frac{v}{u + v})\) the expected frequency of trait 0 when inertia no longer plays a role in its evolution (that is, when an infinite amount of time has passed).

The obvious limitation to this approach is that the values of the parameters \( u, v, \) and \( t \) are unknown for most traits. In such cases, one possibility is to implement the method we now describe.

### Table 2.1

<table>
<thead>
<tr>
<th>( u \leq v )</th>
<th>( u &gt; v )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural selection and inertia</td>
<td>Inertia only</td>
</tr>
<tr>
<td>Natural selection only</td>
<td>Neither</td>
</tr>
</tbody>
</table>

To motivate the logic of the method of controlled comparisons, we first clarify what an adaptive hypothesis predicts about trait evolution.

Consider again the relation of fur length to ambient temperature. One might expect the adaptive hypothesis to assert that longer fur evolved if the environment got colder, and that shorter fur evolved if the environment got warmer. This way of construing the adaptive hypothesis is as natural as it is mistaken. Imagine that the optimal relationship between fur length and temperature is represented by the solid line in Figure 2.1. If natural selection moves species toward this relationship, then it is false that whenever the temperature got colder, longer fur evolved. This happens in lineage A. However, for lineage B, after its environment got colder, natural selection caused shorter fur to evolve. Similarly, it is false that whenever the temperature got warmer, shorter fur evolved. This happened in lineage C, but not in lineage D.
As a result of these discrepancies, we interpret adaptive hypotheses as making a claim only about the direction of change that occurs within a lineage: Natural selection tends to bring a species's trait value closer to its optimal relationship with another trait (or with the environment). This does not imply that observed species will necessarily have trait values that are optimal or even close to being so.

In most comparative analyses, data are typically limited to extant species, even if one has a phylogeny. How can one use such data to test an adaptive hypothesis that describes what occurred in the past? Consider an ancestral species A with traits X and Y that gives rise to two descendants D₁ and D₂, as shown in Figure 2.2. During speciation, species D₁ has its X value reduced and species D₂ has its X value increased; both then evolve toward a new optimal Y value. If D₁ evolves a greater distance toward the optimal trait association than does D₂, the X and Y values of these species will exhibit a positive association. This difference in behavior between the two lineages is expected if trait Y is under directional selection and if the fitness of a given phenotype is a monotonically decreasing function of its distance from the optimal phenotype. In this situation, all other things being equal, the net trait change should be greater for D₁ because the response to selection is proportional to the intensity of selection (e.g., see Falconer 1989).

This is the rationale for believing that an adaptive hypothesis that describes the past makes a prediction about extant species. The prediction is that observed Y values differ in the same direction as their optimal values differ. So, for example, this would imply that the observed Y value of species D₁ in Figure 2.2 is less than that of D₂ because the optimal Y value of D₁ is less than that of D₂. The general prediction that traits differ in the same way as their optima do is not dependent on the optimal relationship between traits being linear.

**SELECTION AND PHYLOGENETIC INERTIA AS PROBABILISTIC INFLUENCES**

An analogy will allow us to describe the probabilistic foundations of the method of controlled comparisons.

What does it mean for smoking and for asbestos exposure to cause cancer? Each entry in Table 2.2 represents the probability of an individual's developing cancer conditional on one of four possible histories of smoking and asbestos exposure. Because a cause raises the probability of its effect when other causal factors are held fixed (Eells 1991), one can conclude that smoking causes cancer if $P_1 > P_2$, and $P_3 > P_4$, and that asbestos causes cancer if $P_1 > P_2$ and $P_5 > P_4$. These
pairs of inequalities are logically independent. It is possible for neither, one, or both to be true.

To avoid being misled by spurious associations, it is essential that one hold fixed one putative cause when assessing the causal role of the other. To see why, suppose that smoking causes cancer but that asbestos does not. (In what follows, $P(x \mid y)$ denotes the probability of event $x$ given event $y$.) Even though asbestos makes no difference to the incidence of cancer, it still can be true that

$$P(\text{Cancer} \mid \text{Asbestos}) > P(\text{Cancer} \mid \text{No Asbestos})$$

This inequality will be true if asbestos exposure and cancer are correlated (suppose that most people either smoke and are exposed to asbestos or don’t smoke and are not exposed). However, this inequality disappears when one holds fixed smoking history, because

$$P(\text{Cancer} \mid \text{Asbestos} & \text{Smokes}) = P(\text{Cancer} \mid \text{No Asbestos} & \text{Smokes})$$

and

$$P(\text{Cancer} \mid \text{Asbestos} \& \text{Does not Smoke}) = P(\text{Cancer} \mid \text{No Asbestos} \& \text{Does not Smoke})$$

This analytical framework carries over to the case of inertia and selection. Consider a set of lineages of equal duration, each of which has the 0, 1 trait we discussed earlier. Each lineage begins with ancestral trait values $X_A$ and $Y_A$ and ends with descendant trait values $X_0$ and $Y_0$. For our example, we formulate the adaptive hypothesis in terms of an optimality model predicting that $Y_0 = 0$ if $X_0 = 0$ and $Y_0 = 1$ if $X_0 = 1$. Here, $Y_0$ denotes the predicted optimal trait for a descendant species.

Table 2.3 describes the probability that a descendant species D has $Y_0 = 1$, conditional on four possible facts. The hypothesis of phylo-

---

Table 2.2. The probability ($P_i$) of cancer, conditional on each of four possible treatments

<table>
<thead>
<tr>
<th></th>
<th>Exposed to asbestos</th>
<th>Not exposed to asbestos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smokes</td>
<td>$P_1$</td>
<td>$P_2$</td>
</tr>
<tr>
<td>Does not smoke</td>
<td>$P_3$</td>
<td>$P_4$</td>
</tr>
</tbody>
</table>

Table 2.3. The probability ($P_i$) that $Y_D = 1$, conditional on each of four possible treatments

<table>
<thead>
<tr>
<th></th>
<th>$Y_0 = 1$</th>
<th>$Y_0 = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_A = 1$</td>
<td>$P_1$</td>
<td>$P_2$</td>
</tr>
<tr>
<td>$Y_A = 0$</td>
<td>$P_3$</td>
<td>$P_4$</td>
</tr>
</tbody>
</table>

---

Genetic inertia asserts that $P_1 > P_2$ and $P_3 > P_4$. This hypothesis says that it is more likely that $Y_D = 1$ if $Y_A = 1$ than if $Y_A = 0$, holding fixed the value of $Y_0$. The hypothesis of selection asserts that $P_1 > P_2$ and $P_3 > P_4$. This hypothesis says that it is more likely that $Y_D = 1$ if $Y_0 = 1$ than if $Y_0 = 0$, holding fixed the value of $Y_A$.

One can also extend this framework to continuously varying traits. As before, the test criteria rely on the assumption that selection or inertia is controlled for if two species have predicted or ancestral trait values that are equal. Such equality is impossible by definition for continuous traits. Accordingly, one must lump similar trait values into discrete classes, ideally in a statistically principled manner perhaps using an information criterion to determine class number and width (Akaike 1974).

So far we have described the adaptive hypothesis as specifying an optimal value of $Y$ for each value of $X$. However, such precision is not necessary. One can use a model that predicts that a species evolves toward a certain region of $Y$ values, given its $X$ value. What is essential is that the model describe the expected direction of trait change.

Implementing the controlled comparison method does not require that one have complete information (e.g., at least one observation for each of the $n^2$ different joint states that exist for a pair of traits each having $n$ states). As in any evolutionary analysis, one can proceed with partial information.

The example we’ve just discussed involves lineages of identical duration. We discuss later how to proceed when the lineages considered have unequal durations.
TESTING THE HYPOTHESES

We have described how each of the hypotheses, adaptive and inertial, generates a set of probabilistic inequalities. As in the cancer example, neither, one, or both of the hypotheses may be supported when one determines whether the corresponding observed frequencies differ significantly in the predicted directions.

An inertial hypothesis says that trait values of ancestors and descendants should be associated, when one controls for the influence of selection. A selective hypothesis makes a prediction whose testing requires that one control for ancestral trait values. Yet only the trait values of extant species are usually observed. How can either hypothesis be tested? Under many plausible models of trait evolution, related extant species will have non-independent trait values. This means that an extant species can serve as a proxy for another species’ ancestor when adaptive and inertial hypotheses are tested. What exactly do these hypotheses predict? Suppose that the adaptive hypothesis predicts that a discretely varying trait Y evolves toward an optimal relationship with trait X. Extant species are formed into pairs of relatives, $D_1-D_2$, $D_3-D_4$, and so on. The inertial hypothesis says that if $D_1$ and $D_2$ have the same optimal Y value, then $D_3$ and $D_4$’s observed Y values should differ in the same direction as $D_1$ and $D_2$’s. The adaptive hypothesis says that pairs of relatives such as $D_1$ and $D_2$ should have Y values that differ in the same direction as their optimal Y values. This formulation underscores the point that the method of controlled comparisons resembles the sign method proposed by Burt (1989) in its use of pairs of related species to evaluate adaptive hypotheses. However, the method of controlled comparisons also allows one to test inertial hypotheses, something that is not possible with the sign test.

Of course, a given species has many extant relatives. Accordingly, there is no uniquely correct pairing of extant species that one must consider when the hypothesis of inertia is tested. Consider a fully resolved “binary” phylogeny such as that shown in Figure 2.3. The pairs of species $D_1-D_2$, $D_3-D_4$, $D_5-D_6$, and $D_7-D_8$ are appropriate when one is testing the hypothesis that the trait values of ancestors at time $a_1$ influenced the trait values of extant species. However, if one wants to test whether trait values of ancestors at time $a_2$ influenced the trait values of extant species, the pairs $D_1-D_5$, $D_2-D_6$, $D_3-D_7$, and $D_4-D_8$ are more appropriate. The problem now is that the lineages connecting the species in one pair overlap with the lineages connecting those in another. This problem of nonindependence is most easily dealt with by using pairs with nonoverlapping lineages.

Most phylogenies are complex enough that any rule for pairing species may result in pairs that differ with respect to the recency of their most recent common ancestors. In such cases, acceptance of the inertial hypothesis might reflect the fact that inertia played a significant role for species pairs that diverged recently but not for those that diverged long ago. It is also possible that rejection of the inertial hypothesis could reflect heterogeneity of times since divergence. How species pairs are formed may also affect the acceptance and rejection of adaptive hypotheses. These points are a reminder that most phylogenies can generate multiple adaptive and inertial hypotheses; as a result, talk of testing for “the” effect of selection or inertia is ambiguous. Instead, one may often need to test a set of contingent hypotheses about these influences.

The use of extant species pairs to test hypotheses of ancestral influence makes sense when one considers how it is analogous to the use of sibs to estimate trait heritability. One can determine narrow-sense heritability by analyzing offspring alone (Falconer 1989). If sibs are more similar in height than two randomly selected individuals in the population, this is evidence that height is heritable; this inference of ancestral influence does not depend on measuring or inferring the heights of parents. By analogy, there is no need to use real or inferred ancestral trait values if one wishes to test the hypothesis of phyloge-
netic inertia. Of course, if one is very confident about the trait values assigned to ancestors, one may want to use those values to test an inertial hypothesis. However, the problematic nature of reconstructing ancestral trait values (see Cunningham et al. 1998) makes it fortunate, in our opinion, that it is not necessary in order to use the method of controlled comparisons.

EVOlUTIONARY EXPLANATION AND NULL HYPOTHESES

We noted at the beginning of the chapter that inertia is commonly regarded as a null hypothesis; as Lewin says, inertia should take precedence over selection as an explanation of why land vertebrates have four limbs, even though both hypotheses are able to explain the observations. We believe that this attitude is mistaken. A causal hypothesis asserts that some factor makes a difference in the probability of an effect variable, and so causal hypotheses are never null. Because inertia and selection describe possible causes of trait values, neither of these hypotheses should be regarded as null. We underscore this point by returning to the analogy with the relationship between cancer, smoking, and asbestos exposure. The conclusion that either environmental factor causes cancer requires the marshaling of evidence, and obviously neither takes precedence over the other.

This point bears on methods such as those of Cheverud et al. (1985) in which one determines the degree to which a species’ trait value is predicted by the trait values of its relatives. High predictability is said to reflect a strong effect of inertia. Residual differences are said to represent the effect of some other process, possibly natural selection. Such an analysis gives precedence to inertial hypotheses over adaptive hypotheses, in our view illegitimately. If relatives have similar trait values, it is impossible to say whether it is due to inertia until one knows how similar the hypothesis of natural selection predicts they should be.

FELSENSTEIN’S METHOD OF INDEPENDENT CONTRASTS

Having described our proposal for testing adaptive and inertial hypotheses—the method of controlled comparisons—we now want to explain how it is related to Felsenstein’s (1985) method of independent contrasts.

As noted earlier, many models of trait evolution imply that the traits of related species are not independent. However, most standard statistical tests require that the data be independent and identically distributed. Felsenstein’s (1985) solution to this problem is based on the insight that when trait evolution is modeled as a Brownian motion process, the difference between two species’ trait values is independent of the difference between the trait values of other species pairs, if the branches of the phylogeny linking the species in one pair to their most recent common ancestor do not overlap with the branches linking species in other pairs to their most recent common ancestors. After standardization to a common distribution, such differences are called independent contrasts. Independence also requires the assumption that there are no common ecological influences affecting the trait in related species.

The important features of contrast calculation can be explained by reference to Figure 2.4. For the two species in the left-hand clade, the contrast with respect to trait $X$ is $\frac{2}{\sqrt{2s^2_{11}}}$, where $2$ is the difference in the species’ trait values, and $s^2_{11}$ is the amount of variance that accumulates in trait $X$ per unit time. For trait $Y$, the contrast for these species is $\frac{4}{\sqrt{2s^2_{11}}}$, where $s^2_{11}$ is the amount of variance that accumulates in trait $Y$ per unit time. For the pair of species in the right-hand clade, we obtain the contrasts $\frac{6}{\sqrt{2s^2_{11}}}$ and $\frac{8}{\sqrt{2s^2_{11}}}$. One also calculates contrasts whose numerators are differences between the average values in the two clades; the resulting $X$ and $Y$ contrasts are $\frac{10}{\sqrt{2s^2_{11}}}$ and $\frac{12}{\sqrt{2s^2_{11}}}$, respectively. Thus, the four species yield three pairs of independent, identically distributed contrasts. One can now test whether a predicted adaptive association is present between the two traits by calculating, say, the correlation between the $X$ and $Y$ contrasts.

This method controls for the possible effect of inertia when testing an adaptive hypothesis (but see Price 1997). This is not the same thing as testing for inertia. To understand the difference, return to our analogy with smoking. Suppose one wants to test the hypothesis that smoking causes cancer, but one is concerned that the sex of
Figure 2.4. A phylogeny that yields three pairs of independent contrasts. $X_D$ and $Y_D$ denote observed traits of extant species. $t_1$ measures elapsed time from the first bifurcation to the bottom of the phylogeny; $t_2$ measures elapsed time from the present back to the first bifurcation.

an individual also may make a difference in susceptibility to cancer. The solution is to ensure that one examines a group of smokers and a group of nonsmokers that have the same mix of males and females. Because the groups are the same in this respect, any difference in the incidence of cancer cannot be due to sex. This design controls for the influence of sex but does not test whether it actually affects susceptibility to cancer.

Felsenstein's assumption that trait evolution can be described by a Brownian motion process is plausible when natural selection randomly changes its direction or when genetic drift occurs (although a scale transformation is required). Brownian motion is not a good model of trait evolution if a trait evolves in the direction of a fixed optimum (such as that shown in Figure 2.2). In general, a model used in testing a hypothesis should be neutral on the question of whether the hypothesis under test is true or false (unless the model is independently justified). One possibility is to derive contrasts using another model that is more consistent with an expected evolutionary dynamic. A natural choice when examining traits evolving toward fixed optima is to assume that an Ornstein-Uhlenbeck process governs trait change. However, we interpret the simulations performed by Díaz-Uriarte and Garland (1996) as indicating that the method of independent contrasts performs poorly in this circumstance.

Can the Method of Independent Contrasts Be Used to Test for Phylogenetic Inertia?

Is there a circumstance in which a method such as Felsenstein's that controls for the effect of a variable also allows one to test for it? Consider again our cancer example. Suppose there is an association between smoking and cancer when one does not control for gender and that this association disappears when gender is controlled for. In this “yes/no” situation, one has not only controlled for the influence of gender but also tested for it. The evidence implies that gender does influence susceptibility to cancer.

This suggests that one could deploy a two-part test for the influence of inertia, by asking whether trait values are associated and whether their contrasts are associated. By analogy with the cancer example, in the “yes/no” case, when traits are associated but their contrasts are not, one could claim to have provided evidence for inertia. Unfortunately, this approach does not work in general. It is straightforward to construct “yes/no” cases that the method of controlled comparisons indicates do not provide evidence for inertia. (In the case of the 0.1 trait we discussed earlier, the lack of evidence for inertia and evidence for selection would occur when $P_1 = P_3 > P_2 = P_4$. See also Table 2.3.)

Similarly, in the “yes/yes” case, when traits are associated and their contrasts are as well, this does not imply that inertia has played no role in trait evolution. It is straightforward to construct “yes/yes” cases that the method of controlled comparisons indicates provide evidence for inertia.

Both of these discrepancies depend on a potentially arbitrary classification of continuously varying traits (assumed by Felsenstein's method) into discrete classes (which the method of controlled comparisons uses). Nonetheless, their existence provides a cautionary tale, especially to those who would conclude that a “yes/yes” case implies that inertia has not played a significant role in trait evolution (see Ricklefs and Starck 1996 for related discussion).

These discrepancies follow from our general point earlier that evidence for either natural selection or inertia does not necessarily imply anything about the presence or absence of the other evolutionary force.

We conclude that the method of independent contrasts does not provide a general procedure for testing the hypothesis of phyloge-
ngetic inertia. This conclusion also applies to the sign test proposed by Burt (1989).

**CONFLICTS BETWEEN THE METHOD OF INDEPENDENT CONTRASTS AND THE METHOD OF CONTROLLED COMPARISONS IN THE TEST OF ADAPTIVE HYPOTHESES**

The method of independent contrasts succeeds in transforming a data set in which there is nonindependence into a data set in which there is independence. In this way, one has controlled for phylogenetic inertia. However, this does not entail that contrast data should be generally used to test adaptive hypotheses (see also Price 1997). There are several reasons for this. First, the model of natural selection most readily compatible with the method’s technique for calculating contrasts is in conflict with the model of selection underlying most adaptive hypotheses. An adaptive optimum undergoing a random walk over time is consistent with the Brownian motion assumption, but we think this is a conception of natural selection that few users of this method would regard as generally justified. Most users have in mind deterministic models of natural selection. A few consider stochastic models. In either case, the predictions that are almost always considered of interest are fixed interior points or intermediate optima, such as might occur when there is stabilizing selection. Because biologists often tend to look first and only for a significant association between the contrasts for a data set in order to support a claim about adaptive association, the conflict between the model and user assumptions about the nature of selection tends to remain hidden. Second, the method relies inherently on the assumptions that traits vary continuously and that trait dynamics are independent of trait value. If a trait is assumed to be continuously distributed, trait values cannot be shared; they must change. Although common in biology, these assumptions about trait change and its rules are in conflict with what many people regard as essential aspects of the biology that motivates many claims about inertia: trait identity (as in Lewin’s claim about vertebrate limb number) or near identity and the fact that some trait values are, in effect, fixed, whereas others are plastic. Ideally, the biology many have in mind in this context motivates a model in which there are a denumerable number of states for a trait, trait values can be shared, and the rules governing trait change are determined in part by present trait value. Our general point is that

the independent contrast method, although well motivated, relies on assumptions as to the nature of traits and the form of natural selection that are quite restricted in their applicability. When these assumptions are met, it is an excellent approach. In other circumstances, other approaches are required. One alternative is the method of controlled comparisons, especially inasmuch as it uses only observed trait values, and it requires that the investigator make an explicit decision as to trait categories and about the way in which natural selection acts. There is no necessary reliance on any one assumption as to the nature of traits (such as whether they are discrete or continuous). It also accomplishes the main goal that the method of independent contrasts sets for itself, which is to control for nonindependence in the data induced by phylogenetic relationships. It permits adaptive hypotheses to be tested properly and also allows inertial hypotheses to be tested. For these reasons, we believe that the method of controlled comparisons is preferable. Another preferable alternative is to use the method described by Hansen (1997) to infer what effect, if any, phylogeny has on the approach of a continuous trait to adaptive optima (see also Martins and Hansen 1997).

**SUMMARY**

Natural selection and phylogenetic inertia have often been portrayed as mutually exclusive explanations of trait values or of trait associations. We have described how neither, one, or both of these explanations may apply in any given case. The recognition of the potential importance of ancestral influence on traits has often been accompanied by the idea that an inertial hypothesis takes precedence over an adaptive hypothesis. We reject this idea. Both hypotheses are causal claims, and, to this extent, neither is a null hypothesis.

Just as testing an adaptive hypothesis requires that one control for the possibility of phylogenetic inertia, so a proper test of an inertial hypothesis requires that one control for the possibility of selection. A specific adaptive hypothesis must be used in order to make these controlled comparisons. This will not be welcome news to biologists who regard the effect of ancestors on descendants as a biological first principle that requires no support from data, nor will it be welcomed by those who think that adaptation is so obvious that the testing of adaptive hypotheses is unnecessary.
Our ideas about the testing of adaptive and inertial hypotheses connect with larger questions about the meaning and testability of adaptationism. We have claimed elsewhere that an ensemble of assessments of the local optimality of traits can underwrite an empirical test of adaptationism (Orzack and Sober 1994a, b; 1996). Assessing the role of phylogenetic inertia in a trait's evolution is an important if not always attainable aspect of assessing whether present evidence supports a claim of local optimality for the trait. In this context, we think it is crucial to see that any causal claim about a trait's evolution requires evidence. In the debate over adaptationism, we believe such an attitude would go far toward creating common understanding, especially in comparison with attitudes such as "Optimality models are useless" and "Traits must be adaptive." Such attitudes only impede the search for truth.

ACKNOWLEDGMENTS

We benefited tremendously from the criticism of Thomas Hansen. We thank Dave Wilson for asking a question about adaptation and phylogenies that led us to undertake this project, and Joe Felsenstein and Emilia Martins for important comments. E. S. thanks Brandon Fidelson for assistance and Jim Crow, Carter Denniston, Ted Garland, and Dick Lewontin for discussions. This research was partially supported by National Science Foundation (NSF) grants DEB-9407965 to S.O. and SES-9906979 to S.O. and E.S.

LITERATURE CITED


