Chapter 1

A Likelihood Framework for the Phylogenetic Analysis of Adaptation

DAVID A. BAUM AND MICHAEL J. DONOGHUE

The common core of all studies of adaptation is the assessment of the role of natural selection in character origin and/or maintenance. A range of information seems pertinent to such studies, including phylogeny, ecology, developmental biology, physiology, biomechanics, ethology, and genetics. At present, however, we lack a conceptual framework for integrating these diverse types of data. Our primary mission in this chapter is to develop one such framework based on likelihood ratios. As we hope to show, this likelihood approach to the study of adaptation not only clarifies the hidden assumptions of adaptationist studies but also provides a common language for communication among disciplines. In particular, we believe that this approach will clarify the interrelationship between studies of uniquely evolved characters and the phylogenetic distribution of analogous variation.

This chapter is not about the definition of adaptation. We happen to prefer a definition of adaptation as characters that evolved via natural selection for some specified biological role (Gould and Vrba 1982; Sober 1984). This historical view challenges us to decipher the causes of the fixation of a particular state in a particular ancestral lineage. As difficult as this might be, we think it is valuable to have a definition of adaptation that encourages such investigations rather than one that leads us to restrict our attention to character function in extant organisms. However, even biologists who prefer an ahistorical definition (e.g., Fisher 1985; Reeve and Sherman 1993) are generally still interested in knowing how and why particular traits evolved. Therefore, regardless of one's favorite definition of adaptation, improved methods for inferring the historical action of natural selection may be welcome. Such methods are the subject of this chapter.

PHYLOGENETIC APPROACHES TO THE STUDY OF ADAPTATION

Among those who maintain that phylogeny bears on the study of adaptation, two distinct schools have emerged. Coddington (1994) dubbed them the "homology" and "convergence" approaches. The homology approach is characterized by a focus on a particular character in a particular lineage and the attempt to test hypotheses of adaptation using phylogenetic and functional information. This approach is primarily practiced by systematists and tends to use the parsimony framework described more fully in the next section. The convergence approach, in contrast, is primarily involved in looking at broad phylogenetic patterns of correlation among characters (discrete or continuous) or between characters and environmental variables. The central objective of the convergence approach is to establish whether such correlations exist when one takes into account the inferred phylogeny (e.g., Felsenstein 1985; Harvey and Pagel 1991). The convergence approach has been popular among ecologists and behavioral biologists.

Whereas the homology approach is directly concerned with evaluating the historical action of natural selection on a particular character and lineage, this is not so obviously the case for the convergence approach. After all, it can be argued that detecting phylogenetic correlations is an end in itself, analogous to discovering biological laws (Pagel 1994a). And however strong a correlation may be, it cannot determine whether any individual evolutionary change was or was not caused by natural selection (Wenzel and Carpenter 1994). Nonetheless, there are many researchers involved in the study of broad phylogenetic patterns who seem to be specifically interested in assessing the role that selection played in the origin of characters (e.g., Ridley 1983; Sillén-Tullberg 1988, 1993; Donoghue 1989; Maddison 1990; Maddison and Maddison 1992). Indeed, we suspect that part of the attractiveness of the convergence approach for evolutionary biologists at large is an intuition that observing a general pattern of correlation in $n-1$ cases sheds some light on the adaptive status of the nth case. However, although we may intuit the existence of such a link, a formal treatment has not been achieved and methods of analysis have not been developed. One objective of our chapter is, therefore, to lay the foundation for such a rigorous treatment of information flow between classes of analogous characters and individual cases.

THE PARSIMONY FRAMEWORK

The use of phylogenies in evaluating hypotheses of adaptation was formalized by Greene (1986), Coddington (1988), and Baum and Larson (1991) based on the conceptual framework of Gould and Vrba (1982). The approach that emerged sought predictions of adaptive hypotheses that could be evaluated with phylogenetic information combined with appropriate functional studies. Two predictions were highlighted: (1) that the character enhances performance of the specified biological role relative to the antecedent state and (2) that the character evolved in a lineage whose selective regime was such that enhanced performance of that biological role was favored. The former test of current utility primarily involves functional studies of the character compared with its inferred antecedent state. Here, phylogenetic information helps identify the character's precursor condition and the lineage in which the transition occurred. The second test evaluates historical genesis. It initially involves identifying aspects of the selective regime that determine whether improved performance of the biological role is selectively advantageous. These aspects of the selective regime are then mapped onto the phylogeny using parsimony (in combination, perhaps, with biogeographic and paleoecological information) so as to evaluate whether selection would have favored the derived over the antecedent state on the lineage in which the character is thought to have evolved (Figure 1.1). If both tests are passed, the character is inferred to have evolved via natural selection for the biological role; that is, the adaptive hypothesis is corroborated. If either test is failed, the adaptive hypothesis is refuted.

This phylogenetic approach has been criticized because phylogenies can be difficult to reconstruct accurately (Reeve and Sherman 1993; Frumhoff and Reeve 1994). Even if we have a correct phylogeny, the use of parsimony to reconstruct ancestral states is fallible, especially if we make incorrect assumptions about the rate of character evolution (Frumhoff and Reeve 1994; for related discussion, see Maddison 1995; Schultz et al. 1996; Schluter et al. 1997; Ree and Donoghue 1998). Furthermore, pleiotropy and epistasis may make it difficult to interpret phylogenetic patterns without careful functional and genetic studies (Lauder et al. 1993; Leroi et al. 1994). However, as formulated, these criticisms of the parsimony approach are weak because any scientific inference will fail if the underlying assumptions are not met. In this case the application of the parsimony

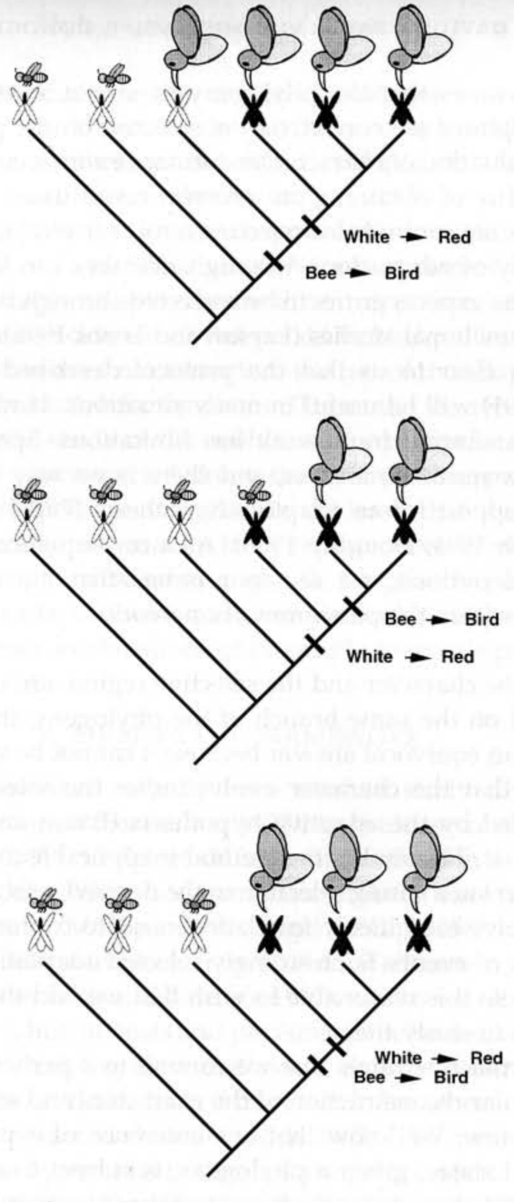

Figure 1.1. A hypothetical example of the parsimony approach to the study of adaptation (see Baum and Larson 1991). A phylogeny is shown for a plant group, some of which have the ancestral condition, white petals, whereas others have the derived condition, red petals (shown in gray). Parsimony is used to locate the branch on which petal color changed. To test the hypothesis that red petals evolved due to natural selection for the biological role of attracting pollinating birds, we first must show that red petals perform better than white petals in bird-pollinated plants. If this is shown, and if parsimony reconstruction of the pollination mechanism (the selective regime as defined by Baum and Larson 1991) implies that red petals evolved on a bird-pollinated lineage (top panel), then the adaptive hypothesis is supported. In contrast, if red petals evolved in a bee-pollinated lineage (middle panel), then the character cannot have evolved via natural selection for bird attraction and, therefore, it is not an adaptation (it is an exaptation sensu Gould and Vrba 1982). If bird pollination and red petals evolve on the same branch (bottom panel), the result is equivocal; the adaptive hypothesis is neither supported nor rejected.

method assumes that the phylogeny is correct, that character and selective regimes are correctly reconstructed on the phylogeny, and that the evaluation of character performance was accurate. Therefore, the possibility of obtaining an incorrect result using the parsimony approach is not grounds for rejecting a role for phylogenetic analysis in the study of adaptation. Although mistakes can be made, in the long run one expects errors to be corrected through further phylogenetic and functional studies (Larson and Losos 1996).

It seems clear to us that the protocol described by Baum and Larson (1991) will be useful in many situations. However, we think that the parsimony framework has limitations. Specifically, it can give only a yes-or-no answer, and there is no way to quantify the *degree* of support for an adaptive hypothesis (Pagel 1994a; Garland and Adolph 1994; Doughty 1996). As a consequence of this lack of a statistical outlook, we see four issues that are not adequately addressed within the parsimony framework.

1. If both the character and the selective regime are inferred to have changed on the same branch of the phylogeny, then the method returns an equivocal answer because it cannot be stated with confidence that the character evolved *after* the selective regime as demanded by the adaptive hypothesis (Baum and Larson 1991; Figure 1.1). This makes the method inapplicable to characters that are under such strong selection in the derived selective regime that they evolve too quickly for cladogenesis to capture the historical ordering of events. Such strongly selected adaptations are of great interest, so it is reasonable to wish that we had the methodological tools to study them.
2. The approach requires that we commit to a particular tree and to a particular reconstruction of the characters and selective regimes on that tree. We know that any inference of a phylogeny or of ancestral states, given a phylogeny, is subject to error. However, the parsimony approach does not take into account such uncertainty, and it does not distinguish between cases in which we are very confident of our historical inferences versus those in which we are unsure.
3. The parsimony approach focuses exclusively on character origin. As discussed earlier, a character that we observe in a living taxon has not only evolved but has also been maintained in at least one lineage through to the present day. However, although understanding the role of selection and other evolutionary forces in this maintenance would seem to be part of a complete adaptationist program, the parsimony methods developed to date are not equipped to shed light on this issue.
4. The approach provides no avenue by which information from elsewhere on the tree can influence our interpretation of a particular case. For example, despite the intuitive importance of observing the repeated coincidence of red petals in bird-pollinated plants, this information does not factor into the test of the adaptive status of red petals in *Lobelia cardinalis*. Thus, the current approach implies that the study of the evolution of a particular character is conceptually disconnected from the study of repeated events (Coddington 1994). We suppose, in contrast, that there is some conduit whereby information from repeated events would influence our strength of belief in a specific adaptive hypothesis and, likewise, that adaptationist studies on particular unique characters would influence explanations of broad phylogenetic patterns.

A LIKELIHOOD FRAMEWORK

Likelihood ratios have recently been used to study directionality in character evolution (Sanderson 1993), to evaluate ancestral states (Schluter et al. 1997), to detect shifts in diversification rate (Sanderson and Donoghue 1994), to detect the correlated evolution of two discrete characters (Milligan 1994; Pagel 1994b, 1998), to compare alternative models of molecular evolution (Felsenstein 1981; Yang et al. 1995; Huelsenbeck and Rannala 1997), and to see whether there has been cospeciation of hosts and parasites (Huelsenbeck et al. 1997). The principle underlying these methods is that the better supported of two hypotheses is that which, if true, would have been more likely to generate the observed data (Edwards 1992). The measure of how much better one hypothesis is than another, the *support*, is given as the natural logarithm of the likelihood ratio (Edwards 1992).

Unlike traditional frequentist statistics (Garland and Adolph 1994; Doughty 1996), likelihoods can be applied to singular observations (Sanderson 1995). As a result, even unique historical events can serve to statistically discriminate between two rival hypotheses. This aspect of likelihoods is possible because the rival hypotheses are evaluated within the context of an explicit probabilistic model. The need to provide such detailed models might be seen as limiting the applicability of likelihood approaches. However, it is commonplace to explore a range of models to determine the area of parameter space within which one hypothesis is favored over the other. This amounts to specifying the background assumptions that need to be met to prefer that hypothesis and is, we think, a great improvement over approaches such as parsimony whose background assumptions are hard to identify (Lewis 1998).

To see how a likelihood framework can be applied to the testing of an adaptive hypothesis, consider the logic of the parsimony approach of Baum and Larson (1991). In this framework, an adaptive hypothesis is supported if a character enhances performance of the specified biological role relative to its antecedent state, and if this enhanced performance was selectively favored at the time of character origination. If both of these things are true, then the parsimony approach assumes that selection for the specified biological role did, indeed, cause character fixation. However, although such an observation strengthens an adaptive hypothesis it is by no means definitive. Even if a derived character has a selective advantage relative to its antecedent state, genetic drift or selection for a different biological role might have been partly or wholly responsible for its fixation. Thus, to evaluate an adaptive hypothesis we must take into account the possibility that the character would have evolved without selection for the hypothesized role. That is, instead of just looking at evidence *for* adaptation, we must also consider evidence in favor of alternative hypotheses. Support for the adaptive hypothesis is, therefore, best captured by the ratio of the likelihood that the character would have evolved via natural selection for the specified biological role versus the likelihood that it would have evolved without such selection.

Consider a situation in which we are studying a particular character (e.g., red petals in *Lobelia cardinalis*), we have a phylogeny in hand, and we believe we know the branch along which red petals evolved. In the likelihood framework we would first seek to evaluate the likelihood of the adaptive hypothesis (L_{adapt}): that red petals would have evolved in the lineage given that there was selection for enhanced attraction of birds. Second, we would seek to evaluate the likelihood of an alternate hypothesis (L_{alt}): for example, that red petals would have evolved in the lineage via genetic drift, without selection for bird attraction. The measure of support for the adaptive hypothesis is the likelihood ratio L_{adapt}/L_{alt}.

A likelihood ratio greater than 1.0 would indicate that the data favored the adaptive hypothesis, whereas a likelihood ratio less than 1.0 would mean that the data argue for the alternative hypothesis. The magnitude of the ratio is a measure of the objective support for the favored hypothesis (Edwards 1992). Various approaches are available to interpret these ratios. Edwards (1992) suggested that a likelihood ratio convincingly favors a hypothesis when the natural logarithm of that ratio exceeds a score of 2.0 (see also Schluter et al. 1997). Alternatively, one can compare the ratio to a null distribution to evaluate its "significance." In some situations one expects log likelihood ratios to follow a χ^2 distribution, with the number of degrees of freedom determined by the difference in the number of free variables in the two models (Felsenstein 1981). In some cases, such as when the two hypotheses are not nested, a χ^2 distribution is not the appropriate null expectation. In those cases the likelihood ratio can be evaluated by Monte Carlo simulation (e.g., Sanderson and Donoghue 1994; Pagel 1994b). It should be noted that there is some question as to how "P-values" should be interpreted in likelihood statistics (Sanderson 1995).

Interestingly, the parsimony approach described earlier can be rephrased as a special case of the likelihood method with certain simplifying assumptions. In effect, it is assumed that L_{alt} is close to zero, meaning that the strength of support for the adaptive hypothesis is basically determined entirely by L_{adapt}. If the character confers a performance advantage for the biological role relative to the antecedent state (i.e., it passes the test of current utility) *and* if the character evolved at a time when enhanced performance of the biological role was selectively favored (i.e., it passes the test of historical genesis), the character is seen as having a high likelihood of evolving (i.e., L_{adapt} takes on an unspecified high value). Given that L_{alt} is assumed to be close to zero, the likelihood ratio in this case would be large and the hypothesis would be supported by the data. If, on the other hand, the character either lacks current utility or evolved at a time when improved performance of the biological role was not favored, then L_{adapt} takes on an unspecified low value, similar in magnitude to L_{alt}. With L_{adapt}/L_{alt} close to 1.0, the null hypothesis cannot be rejected and the adaptive hypothesis is not justified by the data.

The simplifying assumptions built into the parsimony formulation may be reasonable in many circumstances, and, thus, we see no grounds for rejecting that methodology outright. However, the assumption that L_{alt} is low will not apply, for example, if there are rival biological roles that could equally account for the origin of a character. Therefore, the parsimony method might give unwarranted support for the adaptive hypothesis (as noted by Leroi et al. 1994). Genetic drift also must be considered as a potential reason for the evolution of a character, especially if populations at the time of fixation were small. Similarly, L_{adapt} need not always be large – for example, if the performance advantage is minimal. Thus, even if the tests of current utility and historical genesis are passed, an adaptive hypothesis might be only weakly supported by the data. Clearly, it would be desirable to have a framework for evaluating L_{adapt} and L_{alt} directly.

ESTIMATING THE LIKELIHOODS FROM MICROEVOLUTIONARY DATA

Let us imagine that we have identified the branch of a phylogeny along which a character of interest evolved, that is, the derived character is believed to have arisen through mutation and become fixed in the focal lineage. Under the likelihood approach we wish to know the likelihood that the character would have evolved on that branch given either the adaptive or the alternate hypothesis. If we know a lot about the genetics and performance characteristics of the character and about the population biology of the focal lineage, we could, in principle, use a population-genetic model to estimate L_{adapt} and L_{alt}.

One of us (Baum, unpublished) has developed a simple transition matrix approach for cases of randomly mating hermaphroditic diploids with constant selection strength, unigenic characters, and effective population size less than 25. This model assumes that the mutation rate is unaffected by selection and that one knows lineage duration (in number of generations), population size, and the selection coefficient for the character in the relevant selective regime (that of the focal lineage). The probability of the original mutation occurring in any generation is assumed to be equal (i.e., equiprobable prior probabilities of mutation). Given a mutation in generation *i*, the gene frequency in that generation is taken as 1/2N (assuming mutation rate is much lower than 1/2N). The probability of going to fixation by the end of the lineage (P_{fix}) can be obtained using a transition probability matrix taken to the power of the number of remaining generations (Baum, unpublished). One can then obtain the overall likelihood that the character would have evolved by summing P_{fix} across all generations in the focal lineage. Since P_{fix} is influenced by the strength of selection invoked, the likelihood will differ under the adaptive and alternate hypotheses. It should be noted that one does not need to know the actual mutation rate because, if it is assumed to be the same under both hypotheses, it cancels out of the likelihood ratio.

This microevolutionary model was explored for focal lineages of different length (5 to ∞ generations) with different population sizes (5 to 20 diploid individuals) and different selection coefficients (0.025 to 0.4). It was found that the likelihood ratio increases (meaning an increase in the support for the adaptive hypothesis) as the strength of selection increases, as population size increases, and as the duration of the lineage decreases (Baum, unpublished). These results are intuitive. The stronger the estimated selection on the character, the higher is its likelihood of going to fixation in the lineage. The smaller the population size, the greater is the effectiveness of genetic drift relative to selection. The shorter the lineage, the greater the advantage of having selection to drive the character to fixation. The likelihood ratio also increases if one assumes that the character is encoded by more than one gene (Baum, unpublished).

This microevolutionary model is useful for highlighting the factors that affect the magnitude of the likelihood ratio and, hence, the strength of an adaptive hypothesis. However, it would be difficult to obtain good estimates of the input parameters (population size, selection coefficient, etc.). Difficult, but not impossible. Phylogenetic and paleontological data can help one learn about the ecology and population biology of the organisms on the focal lineage. Additionally, one might be able to look at genetic variation in the descendants of the ancestral population to learn something about the effective population size. To determine the genetic architecture and selective consequences of the character in the ancestral population, one would probably study the development, genetics, and performance of the character in extant, descendant populations as compared with suitable surrogates for the antecedent state. In addition, it may be possible to obtain information about character performance by reference to analogous characters occurring in different taxa, or from taxa that are currently experiencing a selective regime similar to that inferred for the ancestral population. Even when we can only narrow the range of possible parameter values, it might still be possible to obtain enough precision to justify support or rejection of the adaptive hypothesis.

Even when we cannot obtain estimates of some of the parameters of the microevolutionary model, we can always make conservative assumptions. For example, assuming that the character is encoded by a single gene and that the lineage in question was infinitely long will reduce the magnitude of support for the adaptive hypothesis. Thus, if one still obtained a convincingly large log likelihood ratio one might conclude that the data at hand strongly supported the adaptive hypothesis.

ESTIMATING LIKELIHOODS FROM ANALOGOUS CHANGES

A category of simplification that warrants special attention is the treatment of analogous characters as replicate experiments. Suppose we are studying the evolution of a particular character, such as the red petals of *Lobelia cardinalis*, and testing the hypothesis that it is an adaptation to a particular biological role, such as attracting pollinating birds. In the absence of microevolutionary information we might be able to extract valuable information from the phylogenetic distribution of analogous characteristics, such as red petals in other angiosperms. For such information to be relevant, one would have to assume "transitivity": that independent occurrences of red petals in other lineages are similar to the case of *Lobelia* in regard to genetic basis and functional consequences, and that the ecology and population structure of the various lineages do not differ greatly from those of the focal case. Under these assumptions each lineage of the phylogeny can be considered an independent opportunity for red petals to evolve. Suppose one were also prepared to assume that the tree is made up of two types of lineages: those in which bird attraction is selectively favored (i.e., the plants are at least partly bird-pollinated) and those in which attracting birds is not beneficial (e.g., those in which birds cannot effectively transfer pollen). If these assumptions apply, then the probability of gaining red petals (P_g) in bird-pollinated (BP) lineages should have a direct relation to the likelihood of red petals evolving given that they are favored by selection for bird attraction ($P_g|BP$). This likelihood should be a reasonable prior expectation of L_{adapt} for *Lobelia cardinalis*. Similarly, P_g in lineages that are not bird-pollinated (NBP) should be a reasonable predictor of L_{alt}.

Pagel (1994b, 1998) and Milligan (1994) have developed a maximum-likelihood method for estimating the probabilities of change based on a tree topology and a set of tip values for both the independent variable (pollination mode) and the dependent variable (petal color). In this method one uses a Markov model to estimate the rates at which lineages gain or lose red petals and bird pollination. In the unconstrained model there are eight distinct rate parameters: four for the gain and loss of red petals with and without bird pollination, and four for the rate of gain and loss of bird pollination with and without red petals. The set of eight parameters that jointly maximize the likelihood of obtaining the tip-states, given the tree, the branch-lengths, and the Markovian model of evolution, is favored. The maximum-likelihood rate parameters can be estimated using the program DISCRETE (Pagel 1994b, 1998). From these rate estimates one can calculate $P_g|BP$ and $P_g|NBP$ for a branch of length t that started without red petals (Pagel 1994b). These transition probabilities should be good a priori estimates of L_{adapt} and L_{alt}, respectively.

We stress here that much more theoretical work is needed on the relationship between the transition probabilities and the likelihoods. In particular, we are cognizant that the transition probabilities estimated by maximum likelihood are point estimates and that they have some associated error. Indeed, our empirical observation (using DISCRETE) is that for trees of moderate size (e.g., <100 taxa) the likelihood surface associated with a given rate parameter is quite flat, such that a wide range of rates (and hence transition probabilities) can equally explain the tip data. Thus, there may be a large error associated with the estimates of $P_g|BP$ and $P_g|NBP$, and one needs a mechanism to take this fact into account. One approach suggested by Pagel (1994b, 1998) is to see whether forcing equality of the rate of gaining red petals in the presence or absence of bird pollination results in a significant decrease in the overall likelihood of obtaining the tip data. Significance in this case can be estimated based either on a χ^2 distribution or a Monte Carlo simulation (Pagel 1994b). One might consider the adaptive hypothesis corroborated only if constraining rate-equality causes a significant decline in the likelihood of obtaining the data. However, this test forces equalities of the rate parameters and not the probabilities of changes in state ($P_g|BP$ and $P_g|NBP$), which are influenced by all eight rates. Yet it is the probabilities of change that seem most relevant to the adaptive hypothesis and the estimation of L_{adapt}/L_{alt} for the specific case. Furthermore, this is a nondirectional test for nonindependence, and, thus, a positive result could arise if $P_g \mid BP$ were *less* than $P_g \mid NBP$.

Given that we have suggested two means of obtaining the likelihood ratio, L_{adapt}/L_{alt} – one based on population genetics and the other based on a broad phylogenetic pattern – it behooves us to briefly comment on how you would proceed if both estimates were available in a given case. The simplest means to integrate such distinct sources of evidence would be to treat them as having equal weight in evaluating support for the adaptive hypothesis. Thus, you could simply add the log likelihoods (see Edwards 1992). Alternatively, one could attempt to develop a model that included population-level information for every lineage of the broader phylogeny. Yet another approach would be to use the data from the broader tree to suggest limits for some parameters in the microevolutionary model or, conversely, to use microevolutionary data to set boundaries on the transition probabilities allowed for the broader comparative analysis.

ADVANTAGES OF A LIKELIHOOD FRAMEWORK

Earlier, we briefly described four limitations of parsimony-based approaches to the phylogenetic analysis of adaptation. As we intimated, a likelihood framework has the potential to deal with each of these limitations. We must stress, however, that further work is needed before this approach will be generally applicable. Our intention here is simply to provide a sketch of how the likelihood framework might accommodate these issues and not to actually implement the relevant methods.

Apparent Simultaneity of Change in a Character and a Selective Regime

The parsimony approach is confounded when a change in selective regime and a change in the character are mapped to the same internal branch (Donoghue 1989; Baum and Larson 1991). This is because it depends critically on ordering the two events using parsimony (perhaps complemented by paleontological information; see Greene 1986). In contrast, the microevolutionary likelihood approach does not depend so much on ordering events but rather considers all possible paths by which a character could become fixed by the end of the lineage. Given a suitable distribution for the probabilities of a shift in selective regime at each generation of the focal lineage, the simplest being the equiprobable distribution, one can calculate the likelihood of fixation over all possible scenarios under both the adaptive and the nonadaptive hypotheses. Intuitively, one expects a large contribution to L_{adapt} to come from scenarios involving a change in the selective regime early in the branch (Figure 1.2), because the chance of the character becoming fixed is higher in the derived than in the ancestral selective regime. If the likelihood of character fixation is high enough under the derived selective regime (relative to the

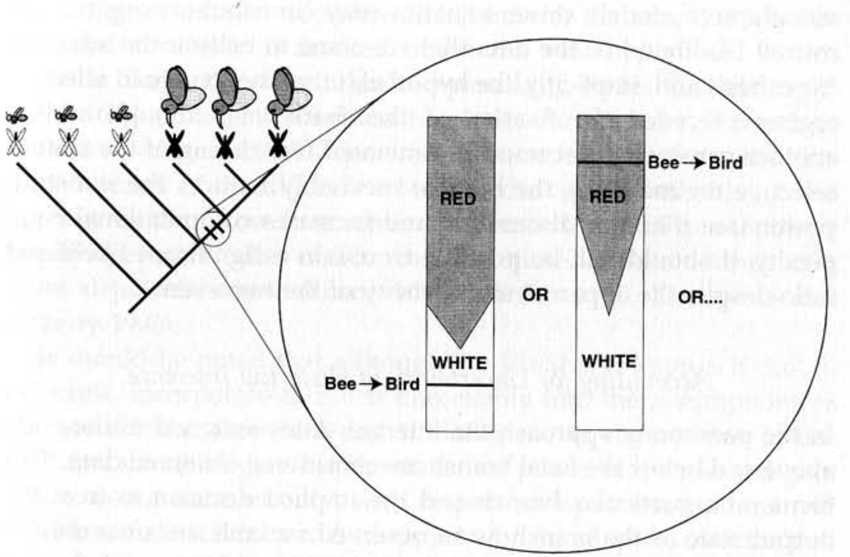

Figure 1.2. The likelihood approach in cases when a parsimony reconstruction of a character places it on the same branch as a shift in some key aspect of the selective regime. Focusing on the lineage on which the character evolved, one asks what is the likelihood that red petals would have evolved under the adaptive hypothesis (selection for bird attraction) versus under the null hypothesis (drift or selection for another biological role). The likelihood approach would consider all possible scenarios for the time of switch in selective regime. For example, the reconstruction on the left shows a shift in pollination type early and then selection rapidly leading to the evolution of red petals. In the second case, red petals went to fixation before the shift to bird pollination and, hence, their evolution did not involve selection for bird attraction. All other scenarios would be considered and would be summed to give the overall likelihood score. Because scenarios involving an early shift in selective regime can contribute a large amount of the overall likelihood under the adaptive hypothesis, L_{adapt} could be found to be significantly higher than L_{alt} even though the regime shift is traced to the same branch as the character's origin.

nonadaptive model), these scenarios may contribute enough to the overall likelihood of the data that we come to believe the adaptive hypothesis and, implicitly, the hypothesis that the change in selective regime preceded the fixation of the character. Although adding another parameter that must be estimated (the timing of the shift in selective regime along the branch) inevitably reduces the statistical power (see the later discussion) and increases computational complexity, it should still be possible to obtain a significant likelihood ratio despite the apparent simultaneity of the two events.

Accounting for Uncertainty in Historical Inference

In the parsimony approach the internal states assigned to the node above and below the focal branch are considered observed data. This focus on a particular branch and the implicit decision to treat the output state of the branch as an observed variable are unavoidable under parsimony. However, within the likelihood framework there is flexibility in what we choose to accept as observed versus what we choose to treat as unknown. When we described the microevolutionary approach to evaluating likelihoods, the ancestral states were taken as being observed such that the test asked whether, given an input of state 0 for the branch in question (plus the model of change), the likelihood that the output state would be 1 is significantly higher under the adaptive than under the alternative hypothesis. However, it should be possible to use a microevolutionary model but to take only the tip states as known. In this case one would calculate the likelihood of obtaining those states under the adaptive and alternative models by summing across all possible sets of ancestral states (appropriately weighted by their likelihood of occurring under the chosen model of evolution). This test would then come closer to the second method we described, which was based on the distribution of tip states, tree topology, and branch lengths and on the transitivity assumptions.

The flexibility of the likelihood approach is not limited to ancestral state reconstructions. In principle, it can accommodate other elements of uncertainty, such as doubt over the tree-topology and/or branch-lengths. In each case, L_{adapt} and L_{alt} are based on the maximum-likelihood score across all plausible values of the unknown variable. For example, imagine that there is one uncertain node in a tree and that, of two possible resolutions, one is favored (i.e., suggests a higher likelihood score) under the adaptive hypothesis, and the other is favored under the null hypothesis. In this case the likelihood ratio might compare L_{adapt} under the first topology with L_{alt} under the second topology. Furthermore, if the trees were derived from DNA sequence data one could also factor the likelihood scores of the different trees into the calculation of L_{adapt} and L_{alt}. The issue of how to accommodate uncertainty in tree-topology in phylogenetic studies of character evolution has been addressed in more detail elsewhere (Losos 1994; Martins 1996; Donoghue and Ackerly 1996).

It should be noted that although the likelihood approach can, in principle, incorporate as much uncertainty into the assumptions as one would like, there is a cost to doing so. As the number of unspecified variables goes up, the statistical power of the method goes down. Being noncommittal with respect to input assumptions will therefore make it harder to find support for the adaptive over the null hypothesis (cf. Sanderson and Donoghue 1995; Huelsenbeck and Rannala 1997). On the other hand, if one is too bold, and if some of the assumptions are actually wrong, then the adaptive hypothesis may be falsely supported. So how should one decide how bold to be?

The answer to this question raises some complex philosophical issues that we cannot resolve here. Taking the noncommittal approach may be seen as more objective, because it takes less for granted. On the other hand, given that we are forced at some level to take *something* as known, the more we put in the known category, the greater is our ability to learn something new. Under the likelihood approach the assumptions made are shaped by how much you think you know at the outset. If you are quite confident that you have correctly identified the branch on which a change occurred, then this should be noted explicitly and used as an assumption in testing the adaptive hypothesis. If you have doubts about ancestral-state reconstructions, then you can choose to accept only the tip-states and the tree as known, and if you are uncertain about the phylogeny you can consider a range of plausible trees (Losos 1994; Donoghue and Ackerly 1996).

Taking Account of Selective Maintenance

The standard parsimony approach to the study of adaptation focuses only on character origin, downplaying selective maintenance (Reeve and Sherman 1993; also see Westoby et al. 1995). Natural selection can be involved in the maintenance of a character when relevant variation arises through mutation but the character persists in a lineage through continued elimination of the alternatives (Ackerly and Donoghue 1995). Fortunately, the likelihood approach can be extended to distinguish maintenance from origin. In the case of selective maintenance, the likelihood ratio depends on a different adaptive hypothesis, namely that the character was maintained in the lineages in which it now occurs due to the character's performance of the biological role. The alternate hypothesis is that the character would have persisted anyway – for example, due to developmental or genetic constraints. To put likelihoods on each of these hypotheses, one might attempt to use a microevolutionary model. However, in this case it would need to take into account the set of other characters that could potentially evolve, something that would be very difficult. Thus, from a practical standpoint, the analysis of adaptive maintenance using a microevolutionary model is even more difficult than that of selective origin.

Even when a microevolutionary modeling approach is not possible, one might be able to look at the distribution of analogous changes. The distribution of gains and losses of the character across a broader tree can provide evidence for or against the adaptive maintenance hypothesis. For example, if testing the hypothesis that red petals were maintained by selection for bird attraction, one would see whether the likelihood of losing red petals is significantly lower on branches of the tree that are bird-pollinated (see Pagel 1994b; Milligan 1994).

Integrating Information from Elsewhere on the Tree

We mentioned earlier that the distribution of characters across a tree can bear on our strength of belief in a specific hypothesis of adaptation. The idea is that if the transitivity assumptions apply, each independent origin of an analogous character represents a replicate "natural experiment" (Doughty 1996). As a result, the frequency with which analogous characters evolve in the presence versus in the absence of the relevant selective regime tells us something about the magnitude of L_{adapt} versus L_{alt} for the specific case. This aspect of the likelihood approach is, we think, a great improvement over existing approaches. However, as mentioned earlier, further work is needed to clarify the statistical basis for such information flow.

CONCLUSIONS

It is important to note that the approach we have outlined is unlike either of the two approaches described by Coddington (1994). It differs from the homology approach in being statistically oriented and allowing the possibility of using broad phylogenetic patterns. It differs from the convergence approach in seeing phylogenetic correlations not as ends in themselves but rather as a means to learn about specific historical events (see Maddison and Maddison 1992, 28–30, for more on this distinction). It seems that in the debate over the relative merits of homology versus convergence approaches, other possibilities, such as the approach we have proposed here, have been overlooked. Our likelihood approach can provide a flexible framework for studying character evolution without abandoning statistical rigor. It permits the analysis of unique historical events, while also taking into account information from analogous characters. We think, therefore, that with further development this approach may offer a useful alternative framework for the study of evolutionary causation.

ACKNOWLEDGMENTS

We are grateful to Steven Orzack and Elliott Sober for inviting us to participate and for forcing us to finish the manuscript. We also thank Steven Orzack and three anonymous reviewers for helpful comments on earlier drafts of this paper. David Ackerly, Allan Larson, Mark Pagel, Rick Ree, and Mike Sanderson provided helpful discussion.

LITERATURE CITED

Ackerly, D. D., and M. J. Donoghue. 1995. Phylogeny and ecology reconsidered. *Journal of Ecology* 83: 730–733.

Baum, D. A., and A. Larson. 1991. Adaptation reviewed: A phylogenetic methodology for studying character macroevolution. *Systematic Zoology* 40: 1–18.

Coddington, J. A. 1988. Cladistic tests of adaptational hypotheses. *Cladistics* 4: 3–22.

Coddington, J. A. 1994. The roles of homology and convergence in studies of adaptation. In *Phylogenetics and ecology*, ed. P. Eggleton and R. Vane-Wright, 53–78. London: Linnean Society of London, Academic Press.

Donoghue, M. J. 1989. Phylogenies and the analysis of evolutionary sequences, with examples from seed plants. *Evolution* 43: 1137–1156.

Donoghue, M. J., and D. D. Ackerly. 1996. Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philosophical Transactions of the Royal Society of London* B 351: 1241–1249.

Doughty, P. 1996. Statistical analysis of natural experiments in evolutionary biology: Comments on recent criticisms of the use of comparative methods to study adaptation. *American Naturalist* 148: 943–956.

Edwards, A. W. F. 1992. *Likelihood: Expanded edition*. Baltimore, MD: Johns Hopkins University Press.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.

Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125: 1–15.

Fisher, D. C. 1985. Evolutionary morphology: Beyond the analogous, the anecdotal and the ad hoc. *Paleobiology* 11: 120–138.

Frumhoff, P. C., and H. K. Reeve. 1994. Using phylogenies to test hypotheses of adaptation: A critique of some current proposals. *Evolution* 48: 172–180.

Garland, T., Jr., and S. C. Adolph. 1994. Why not to do two-species comparative studies: Limitations on inferring adaptation. *Physiological Zoology* 67: 797–828.

Gould, S. J., and E. S. Vrba. 1982. Exaptation: A missing term in the science of form. *Paleobiology* 8: 4–15.

Greene, H. W. 1986. Diet and arboreality in the emerald monitor, *Varanus prasinus*, with comments on the study of adaptation. *Fieldiana Zoology* 31: 1–12.

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

Huelsenbeck, J. P., and B. Rannala. 1997. Phylogenetic methods come of age: Testing hypotheses in an evolutionary context. *Science* 276: 227–232.

Huelsenbeck, J. P., B. Rannala, and Z. Yang. 1997. Statistical tests of host-parasite cospeciation. *Evolution* 51: 410–419.

Larson, A., and J. B. Losos. 1996. Phylogenetic systematics of adaptation. In *Adaptation*, ed. M. R. Rose and G. V. Lauder, 187–220. San Diego, CA: Academic Press.

Lauder, G. V., A. M. Leroi, and M. R. Rose. 1993. Adaptations and history. *Trends in Ecology and Evolution* 8: 294–297.

Leroi, A. M., M. R. Rose, and G. V. Lauder. 1994. What does the comparative method reveal about adaptation? *American Naturalist* 143: 381–402.

Lewis, P. O. 1998. Maximum likelihood as an alternative to parsimony for inferring phylogeny using nucleotide sequence data. In *Molecular systematics of plants*, 2nd edition, ed. P. S. Soltis, D. E. Soltis, and J. J. Doyle, 132–163. New York: Chapman and Hall.

Losos, J. B. 1994. An approach to the analysis of comparative data when a phylogeny is unavailable or incomplete. *Systematic Biology* 43: 117–123.

Maddison, W. P. 1990. A method for testing the correlated evolution of two binary characters: Are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution* 44: 539–557.

Maddison, W. P. 1995. Calculating the probability distributions of ancestral states reconstructed by parsimony on phylogenetic trees. *Systematic Biology* 44: 474–481.

Maddison, W. P., and D. R. Maddison. 1992. *MacClade: Analysis of phylogeny and character evolution*. Version 3. Sunderland, MA: Sinauer Associates.

Martins, E. P. 1996. Conducting phylogenetic comparative studies when the phylogeny is not known. *Evolution* 50: 12–22.

Milligan, B. G. 1994. Estimating evolutionary rates for discrete characters. In *Models in phylogeny reconstruction*, ed. R. W. Scotland, D. J. Siebert, and D. M. Williams, 299–311. Systematics Association special volume No. 52. Oxford: Clarendon Press.

Pagel, M. D. 1994a. The adaptationist wager. In *Phylogenetics and ecology*, ed. P. Eggleton and R. Vane-Wright, 29–51. London: Linnean Society of London, Academic Press.

Pagel, M. D. 1994b. Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London* B 255: 37–45.

Pagel, M. D. 1998. Inferring evolutionary processes from phylogenies. *Zoologica Scripta* 26: 331–348.

Ree, R. H., and M. J. Donoghue. 1998. Step matrices and the interpretation of homoplasy. *Systematic Biology* 47: 582–588.

Reeve, H. K., and P. W. Sherman. 1993. Adaptation and the goals of evolutionary research. *Quarterly Review of Biology* 68: 1–32.

Ridley, M. 1983. *The explanation of organic diversity: The comparative method and adaptations for mating*. New York: Oxford University Press.

Sanderson, M. J. 1993. Reversibility in evolution: A maximum likelihood approach to character gain/loss bias in phylogenies. *Evolution* 47: 236–252.

Sanderson, M. J. 1995. Objections to bootstrapping phylogenies: A critique. *Systematic Biology* 44: 299–320.

Sanderson, M. J., and M. J. Donoghue. 1994. Shifts in diversification rate with the origin of angiosperms. *Science* 264: 1590–1593.

Sanderson, M. J., and M. J. Donoghue. 1995. Reconstructing shifts in diversification rates on phylogenetic trees. *Trends in Ecology and Evolution* 11: 15–20.

Schluter, D., T. Price, A. Ø. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51: 1699–1711.

Schultz, T. R., R. B. Cocroft, and G. A. Churchill. 1996. The reconstruction of ancestral character states. *Evolution* 50: 504–511.

Sillén-Tullberg, B. 1988. Evolution of gregariousness in aposematic butterfly larvae – A phylogenetic analysis. *Evolution* 42: 293–305.

Sillén-Tullberg, B. 1993. The effect of biased inclusion of taxa on the correlation between discrete characters in phylogenetic trees. *Evolution* 47: 1182–1191.

Sober, E. 1984. *The nature of selection: Evolutionary theory in philosophical focus*. Cambridge, MA: MIT Press.

Wenzel, J. W., and J. M. Carpenter. 1994. Comparing methods: Adaptive traits and tests of adaptation. In *Phylogenetics and ecology*, ed. P. Eggleton and R. Vane-Wright, 79–101. London: Linnean Society of London, Academic Press.

Westoby, M., M. R. Leishman, and J. M. Lord. 1995. On misinterpreting the "phylogenetic correction." *Journal of Ecology* 83: 531–534.

Yang, Z., N. Goldman, and A. Friday. 1995. Maximum-likelihood trees from DNA sequences: A peculiar statistical estimation problem. *Systematic Biology* 44: 384–399.