

---

# I.16

---

## Phylogenetics and Comparative Methods

David D. Ackerly

### OUTLINE

1. The role of phylogenetics in ecology
2. Phylogenies and the analysis of trait correlations
3. Phylogenetic signal: Pattern and significance
4. Phylogenetics and community ecology
5. Prospects for the future

The study of ecology frequently draws on comparative observations and experiments that rely on the similarities and differences among species and the correlations among species traits and the environment. In such studies, consideration of the phylogenetic relationships among species provides valuable information for statistical inference and an understanding of evolutionary history underlying present-day ecological patterns. From a statistical perspective, related species do not necessarily provide independent data points for hypothesis tests, due to inheritance of shared characteristics from common ancestors. This similarity can be addressed through a variety of statistical techniques, including the widely used method of phylogenetic independent contrasts. Independent contrasts play a particularly valuable role in the analysis of trait and trait–environment correlations and may point toward alternative interpretations of comparative data. In community ecology, measures of the phylogenetic clustering or spacing of co-occurring species provide a useful tool to test alternative processes underlying community assembly. Co-occurrence of close relatives most likely reflects ecological filtering, in which related species with similar traits share the ability to tolerate local conditions. The reverse pattern of phylogenetic spacing of co-occurring species may reflect a variety of processes, and additional observations of species traits in relation to environment and interacting taxa will be necessary to address underlying processes. Use of comparative methods has increased dramatically with the rapid growth in phylogenetic information and computing power and will continue to play an important role in ecological research.

### GLOSSARY

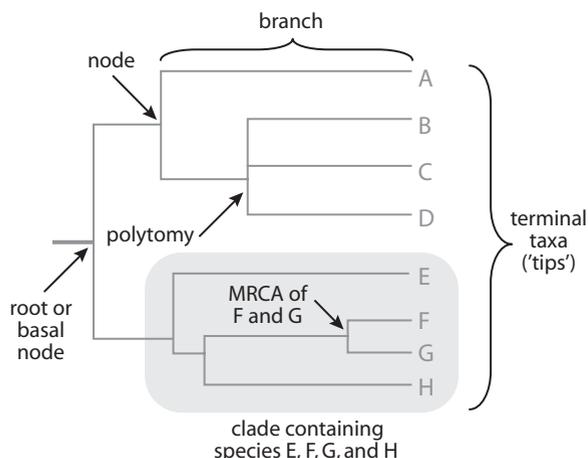
See figure 1 for illustrations of main terms.

**branch lengths.** These may indicate either the number of inferred character changes or a measure of relative or absolute time along any particular branch connecting two nodes. If the molecular data underlying a phylogeny do not violate a molecular clock, a single rate may be imposed such that branch lengths will represent *relative time*, and contemporaneous taxa will be placed at the same distance from the root (i.e., the same age). If a molecular clock is violated, *rate-smoothing methods* have been developed to obtain the best-supported estimate of relative time. Fossils and biogeographic or paleoecological information may then be used to calibrate these branch lengths and convert them to units of *absolute time*. Rate-smoothing and calibration methods are fraught with difficulty, and branch lengths should be treated with caution. (Note that branch lengths may also be set arbitrarily for convenience when one is drawing trees, in which case they have no intrinsic biological meaning.)

**character states.** Phylogenetic trees are reconstructed based on analysis of a matrix of *characters*, where each character can take on one of two or more *states* (binary or multistate, respectively) for each taxon in the group. Phylogenies can be reconstructed from molecular and/or morphological data, although the former are now much more common. Analyses that include morphological data are advantageous as they make it possible to incorporate taxa or fossils for which molecular data are not available.

**lineage.** This refers to a single line of ancestor–descendant relationship, connecting nodes within a phylogeny.

**most recent common ancestor (MRCA).** The MRCA is the most recent node that is shared by any two taxa in a tree.



**Figure 1.** Example of a phylogenetic tree for eight taxa (A–H), illustrating some of the terms in the glossary. This tree is ultrametric, meaning that all terminal taxa are equidistant from the root of the tree.

**phylogenetic distance.** The phylogenetic distance between two nodes or taxa refers to the sum of branch lengths from one tip (or internal node) down to the MRCA and back up to another tip (or node) of a tree. The phylogenetic distance matrix is an  $n \times n$  matrix (for  $n$  taxa) of such distances among all pairs of taxa, with 0s in the diagonal.

**phylogeny.** A phylogeny, or phylogenetic tree, is a branching diagram showing the hierarchy of evolutionary relationships among a group of taxa (extant and/or extinct). Terminal taxa or *tips* are connected by *branches* to internal *nodes* that indicate a hypothesized ancestor. A *clade* includes all of the taxa (extant and extinct) that descend from a node. Phylogenies can be either *rooted* or *unrooted*, where the root represents the hypothesized ancestor of all taxa on the tree.

**polytomy.** This refers to a node with three or more daughter nodes. A *soft polytomy* indicates uncertainty, where the true bifurcating relationships among the daughters are unknown. A *hard polytomy* represents a hypothesis of near simultaneous divergence where the sequence of individual speciation events cannot be meaningfully resolved. Most phylogenetic comparative methods treat polytomies as either hard or soft but do not always make the distinction explicit.

**ultrametric.** An ultrametric tree is one in which all terminal taxa are contemporaneous; more precisely, the sum of the branch lengths from the root to each tip is the same for all tips. Phylogenies of extant taxa will be ultrametric if branch lengths have been adjusted to represent relative or absolute time.

In Great Britain there are 32 indigenous trees[:] of these 19 or more than half . . . have their sexes separated—an enormous proportion compared with the remainder of the British flora: nor is this wholly owing to a chance coincidence in some one Family having many trees & having a tendency to separated sexes: for the 32 trees belong to nine Families, & the trees with separate sexes to five Families.

—Charles Darwin, manuscript for *Natural Selection* (unpublished)

In the quote above, Darwin observes an interesting pattern among plant species of Great Britain. He notes that among trees, the proportion of species that have individuals of separate sexes (as in humans and most vertebrates) is much higher than among the flora as a whole, most of which is composed of shrubs and herbaceous plants. He explained the high frequency of separate sexes as an adaptation to promote cross-fertilization in trees: because trees are large and have many flowers, the chance that an insect would carry pollen from one flower to another of the same individual is quite high. If all the flowers on a tree are of the same sex, these repeated visits by pollinators will not lead to high levels of self-fertilization.

Darwin's observations provide a nice example of what we now call comparative biology, which draws on comparisons of the similarities and differences among species to test ecological and evolutionary hypotheses. In addition, what Darwin recognized intuitively is that a simple count of the number of species exhibiting different characteristics might not be adequate to support his argument. If many of the species are drawn from the same family (that is, closely related in evolutionary terms), they are likely to share many ecological characteristics. Thus, a group with many tree species may also contain many species with separate sexes, reflecting their descent from a common ancestor. But if the evolutionary argument is sound—that trees should evolve separate sexes because of the problem of self-fertilization—then this combination of traits should evolve independently in many different taxonomic groups, and this is indeed what Darwin observed.

Throughout the past 150 years, since the publication of Darwin's *On the Origin of Species*, comparative biology has played a central role in ecology and evolutionary biology. In essence, each species alive today (or in the past) represents the outcome of a long, natural experiment. The results reflect the contemporary ecology of a species—interactions with the abiotic environment and with other forms of life—as well as the cumulative legacy of the past. Evolution works slowly,

and most features are passed down from ancestor to descendant with little change. A penguin appears beautifully adapted to the challenges of surviving and reproducing under the extreme conditions of Antarctic life. But these adaptations must be understood in historical context: penguins are birds, and this experiment in polar living started with very specific initial conditions, including egg-laying, a feathered pelt, forelimbs modified into wings, and so on. Comparative research, placing penguins in the broader context of other birds and viewing them side by side with their closest relatives (loons, albatrosses, petrels, and shearwaters) is critical to an understanding and appreciation of their contemporary ecology and behavior.

In the past 30 years, comparative biology has grown rapidly as a new generation of methods emerged, combining the historical perspective outlined above with the quantitative tools of experimental statistics. The emergence of modern phylogenetics triggered these developments. The word *phylogeny* refers to the evolutionary relationships among a group of organisms, illustrated as a branching tree where the tips (or leaves) may represent individuals, populations, species, or groups of species, and the internal branching points are their common ancestors. The study of phylogenetics has been revolutionized by the combination of molecular biology (providing a trove of data), conceptual advances (the theory of cladistics), and the availability of high-speed computers. Together, these advances have made it possible to infer highly resolved phylogenies for many groups of organisms. With continuous improvements in methods and the availability of data, the tree of life is taking shape and revealing the hierarchy of evolutionary relationships among living (and extinct) organisms.

### 1. THE ROLE OF PHYLOGENETICS IN ECOLOGY

The science of ecology studies the interactions of organisms with their environment and the consequences of these interactions for where species live and how they interact. To address these questions, it is often useful to compare different species, either through observations or experiments. The similarities and differences in how species respond to their environment or interact with each other can provide important ecological insights. When data are gathered on different species, understanding how they are related to each other (i.e., their phylogenetic relationships) contributes valuable information that can affect data analysis and interpretation. In this chapter, I focus on two areas of ecological research where phylogenies play a particularly important role: the analysis of correlations among species traits and environmental conditions (like Dar-

win's example above) and the study of community ecology. In addition, I provide a brief discussion of the concept of *phylogenetic signal*, a general term for the similarity among close relatives.

In the discussion below, it is assumed that a phylogeny is available for each group under consideration. Most phylogenies are based on molecular data, particularly DNA sequences, sometimes combined with morphological or other characteristics. The computational methods used to search for the best-supported phylogeny are continually being improved and are beyond the scope of this chapter. Regardless of the method used, it is important to recognize that every phylogeny is a hypothesis of relationships, and like any scientific hypothesis, it is subject to revision and improvement. Phylogenies may also contain different degrees of uncertainty, both in terms of the topology (the pattern of who is related to whom) and the lengths of the branches, which represent the amount of evolutionary change or the amount of time elapsed between different nodes of the tree. This uncertainty can be incorporated into comparative analyses; in many cases, the results are quite robust across a range of possible alternatives, so strongly supported and fully resolved phylogenies are not a prerequisite for comparative analysis. An overview of some terminology used to describe phylogenies is provided in the Glossary.

### 2. PHYLOGENIES AND THE ANALYSIS OF TRAIT CORRELATIONS

Research in functional ecology, life history strategies, and related areas of ecology often addresses questions of interspecific trait–trait and trait–environment associations, such as: Do mammals with larger body sizes have larger home ranges? Do plant species of open habitats tend to have smaller seeds? How are the traits of invasive species different from those of native species in a community? The answers to these questions help us to understand how species traits influence distribution, abundance, and interactions with other species in a community. They also have important applications in conservation biology, restoration ecology, and the management of invasive species.

A variety of statistical methods can be applied to test hypotheses of trait associations, depending on the type of data available and the nature of the hypothesis. These include correlation, regression, analysis of variance, contingency table analysis, and others. One of the basic assumptions of virtually all statistical tests is that each data point represents an observation that is independent with respect to the underlying null hypothesis. This assumption is not required in order to calculate the various statistics; rather, it is essential to

deriving the statistical significance of the outcome. For conventional statistics, this significance value (or  $p$ -value) represents the probability of observing the data if the underlying null hypothesis is true. When that probability is too low (conventionally, we use a cutoff value of 5%), we reject the null hypothesis and accept that there is a significant effect or relationship. For maximum-likelihood tests, which are playing an increasingly important role in ecology and comparative methods, the assumption of independence is used to assess the likelihood of the best-fit model relative to alternative models or hypotheses, given the observed data.

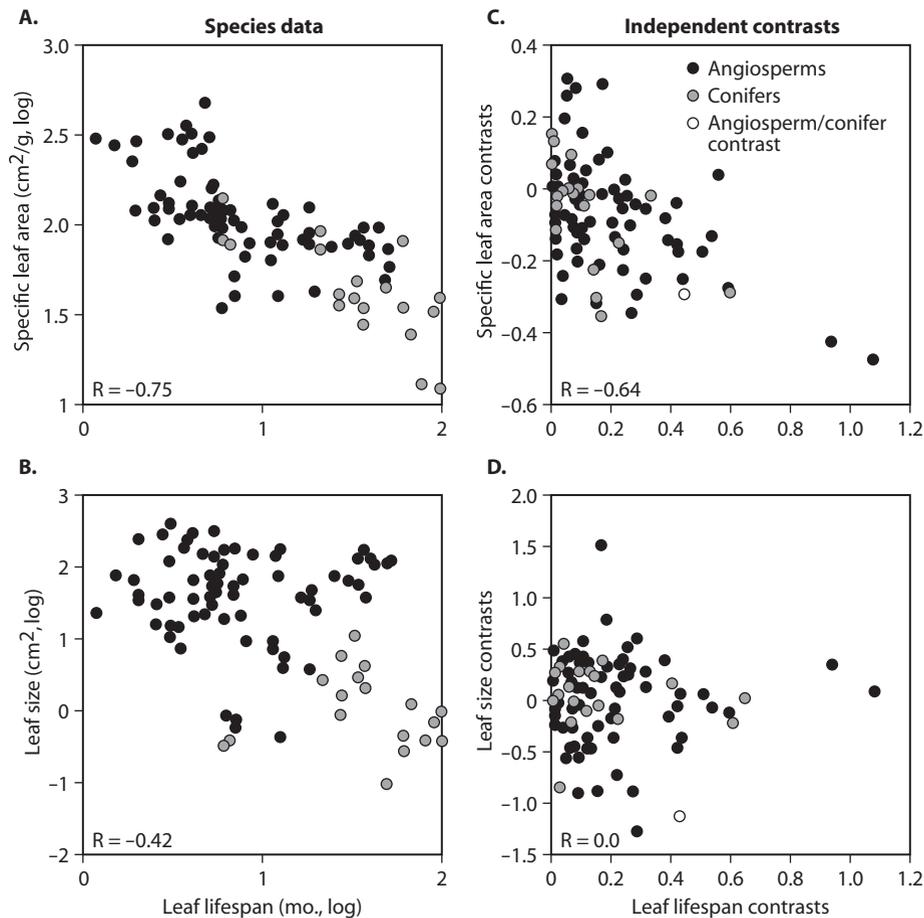
The fundamental argument underlying the development of many comparative methods arises from the observation that related species are ecologically and phenotypically similar to one another. This will not hold for every trait, as instances of rapid divergence and of convergent evolution are widespread and important. But on average, species resemble their close relatives more than they do more distant taxa, and this similarity reflects descent from recent common ancestors. Because of this inherited similarity, it is argued that in statistical terms species do not represent independent data points, violating this basic assumption of significance testing. One can also approach this problem in terms of the underlying historical processes. Trait associations among extant species arise through a historical sequence of correlated changes occurring along each branch of the phylogeny; ideally, we would like to estimate the correlation between these changes to more directly measure evolutionary linkages between the traits. It is now well established that the correlations observed among living species (at the tips of the phylogeny) do not provide a reliable estimate of this historical pattern of correlated evolutionary changes that have occurred along the branches of the phylogeny. Although some researchers are strongly motivated by the statistical arguments, and others more by the historical questions, both perspectives lead one to the use of phylogenetic comparative methods.

The selection of a comparative method to conduct associational analyses depends on the nature of the data and the hypothesis. One of the most common problems is the correlation (a measure of the strength of association) between two traits measured on a continuous scale (e.g., body size or seed size). Correlation coefficients range from 0 for two traits with no association up to 1 for traits that are very tightly linked (−1 if it is a negative association). In 1985, Joseph Felsenstein introduced the method of phylogenetic independent contrasts (often referred to as PICs) to address this question in a phylogenetic context; more than 20 years later, his method remains one of the most robust and

widely used of all comparative methods. The method of independent contrasts rests on the assumption that the evolutionary change in a trait that occurs along each lineage leading up to present-day species represents an independent event with respect to the changes occurring in other branches. Independence, in this context, refers to the statistical notion that the changes are independent manifestations of underlying processes, although the same processes (e.g., natural selection as a result of climate change) may be affecting multiple lineages in a group. If the trait changes that occur in two lineages arising from a common ancestor are independent, then, as Felsenstein demonstrated based on statistical theory, the difference between the trait values of the two descendants will also represent a statistically independent observation. These differences are calculated by subtracting the trait value of one species from the value of its closest relative, and they are referred to as PICs (there is an additional step involving the branch lengths on the phylogeny, which I do not describe here). In addition, Felsenstein showed that one can continue to calculate contrasts at deeper nodes of the tree, based on an iterative process of averaging the trait values at successively deeper nodes. In a fully resolved phylogeny,  $N$  species will be connected by  $N - 1$  common ancestors, so  $N$  trait values measured on the species will provide  $N - 1$  contrasts; these contrasts can be used as the variables in correlation, regression, and multivariate statistical analyses.

A study that I conducted with Peter Reich in 1999 illustrates the application of independent contrasts and how they can impact the analysis of trait associations. We examined correlations among several functional attributes of leaves, including leaf size, leaf lifespan (the length of time a leaf persists on a plant), and specific leaf area (SLA, the ratio of leaf area to leaf dry mass; higher values indicate thinner or less dense leaves). Global studies of leaf function have found that leaves with higher SLA tend to have faster metabolic rates and shorter leaf lifespan, and this strategy is favored in more fertile habitats. The opposite set of traits is observed in leaves with low SLA. In addition, it is sometimes observed that leaves with low SLA and long leaf lifespan are smaller, and small leaves are often viewed as an adaptation to low-water or high-temperature environments. In particular, the needles of conifers (pines, spruces, etc.) are smaller in area and have a longer lifespan than the leaves of most flowering plants.

In a data set of about 100 species, including both conifers and flowering plants, there are negative correlations of leaf lifespan with both SLA and leaf size. However, when we apply independent contrasts, the results change dramatically. The evolutionary correlation



**Figure 2.** Analysis of interspecific correlations among leaf traits, using independent contrasts. Panels A and B show the correlations of leaf lifespan with leaf size and specific leaf area, respectively. Black circles are data for flowering plant species; gray circles are for conifers. The strength of the associations is indicated by the correlation coefficients in the lower left corner of each panel. Panels C and D show the corresponding relationships analyzed with independent contrasts. Black circles are contrasts between nodes

within the flowering plant phylogeny, and gray circles are contrasts among conifers. The white circles represent the contrast at the basal node between the two groups. For convenience, the subtraction at each node is arranged such that the contrast for leaf lifespan is positive, and then the contrast for the other trait is positive or negative, depending on the trait values (subtraction must be in the same direction for both traits). [From Ackerly et al., 2000, *Bioscience*; copyright American Institute of Biological Sciences]

between SLA and leaf lifespan, based on contrasts, is similar to the pattern observed without using independent contrasts. But the evolutionary correlation between leaf lifespan and leaf size is essentially zero (figure 2). Why does this result shift so dramatically? As noted above, most of the correlation observed between leaf lifespan and size results from the marked difference between these traits in conifers and flowering plants, the deepest split in the phylogeny for this group of plants. Independent contrasts capture this shift as a single contrast. The rest of the contrasts, calculated among species of flowering plants or among species of conifers, exhibit no correlation in the shifts

occurring in these two traits. In essence, the pattern observed if each species is treated as an independent data point reflects the influence of a single event deep in the evolution of these groups; when this single event is represented as one data point in the analysis (based on the one contrast), its influence is diminished, and we see that there is not a consistent evolutionary tendency for correlated changes between these two traits. Other lines of evidence are consistent with this result: there is no evidence that leaf lifespan and leaf size are functionally or evolutionary linked to each other, so the result from independent contrasts proves more reliable. We are still left with an important pattern in the

present day: it is true that conifers have small, long-lived needles, which differ on average from the leaves of flowering plants. These differences may be important to understanding the ecological differences between these two groups of plants, but they should not be taken as evidence of an ongoing functional and evolutionary linkage between these traits.

As shown in this example, the method of independent contrasts addresses both the statistical and historical issues associated with the analysis of interspecific trait correlations. The contrasts are statistically independent, so significance values are reliable. The correlation or regression coefficients between the contrasts provide a much more precise measure of the underlying evolutionary pattern compared to a correlation of trait values from present-day species. However, like all statistical methods, independent contrasts invoke key assumptions, and these assumptions have been the source of some controversy. The most important assumption is that trait evolution conforms to a pattern of change known as a constant-variance random walk or Brownian motion. This model assumes that the changes occurring in each unit of time are equally likely to be positive or negative and are drawn from a normal distribution, such that small changes are more likely than large ones. Because these changes accumulate across multiple time steps, the total change along a branch is also expected to be proportional to the length of the branch. On the one hand, simulations have shown that statistical tests based on independent contrasts are quite robust to a variety of deviations from these basic assumptions, particularly if appropriate steps are taken to transform data or branch lengths in advance of analysis. In addition, the Brownian motion model is a reasonable first approximation of a model of evolutionary change based on our knowledge of quantitative genetics and the inheritance of continuous traits. On the other hand, a variety of other models of trait evolution may be considered; under some of these alternatives, species trait values are relatively independent of each other, and independent contrasts (or other comparative methods) do not necessarily provide a reliable measure of historical patterns.

There are several other classes of comparative methods that can be used for questions of trait associations. One of the most important is known as the phylogenetic regression, introduced by Alan Grafen in 1989, or phylogenetic general linear models. These approaches utilize statistical methods in which the user can specify the degree of independence among observations. The phylogeny is used to generate what is known as a variance-covariance matrix, which captures the expected degree of dependence among each

pair of species in a study. This then opens up the full power of linear models, including multifactorial analysis of variance or covariance, with appropriate adjustment of significance tests reflecting the phylogeny. Although this facilitates a much broader range of hypothesis tests, one drawback is that the interpretation of results in terms of underlying historical processes is generally not as straightforward. A related class of methods uses maximum-likelihood approaches to find the best-fit model for a given set of interspecific trait data, given the phylogeny and alternative hypotheses of how the traits may be associated with each other. Maximum-likelihood approaches (and related Bayesian methods) have the general advantage that it is easier to invoke alternative underlying models of trait evolution. Further discussion of these methods, and the issues of branch lengths and evolutionary models, is beyond the scope of this chapter; researchers who will be using contrasts or other methods discussed here are well advised to seek a deeper understanding of these issues.

It is important to note that discrete characters, such as presence/absence of a trait or different states of a morphological character, usually require different approaches. Traditional tests of association for discrete characters involve chi-square or G-tests, based on contingency tables showing the frequency of different pairs of states. Phylogenetic approaches can be used to reconstruct historical transitions from one state to the other and then to test for associations between these transitions or between transitions in one character and the background state of the other character. Maximum-likelihood models, such as the DISCRETE program introduced by Mark Pagel, provide powerful solutions to this problem by testing whether the probabilities of transitions in different characters are associated with each other (see box 1).

---

#### BOX 1. SOFTWARE FOR PHYLOGENETIC COMPARATIVE METHODS

Phylogenetic comparative methods are computationally intensive, and a variety of software packages have been introduced that implement different tests. A few of the most important are briefly summarized here.

**MacClade**, first introduced by David and Wayne Maddison in 1987, set the standard for graphical elegance and ease of use in phylogenetic software. It is primarily used for reconstructing the evolution of discrete characters, based on parsimony methods, and also has limited capabilities for continuous characters.

**Mesquite**, also developed by the Maddisons, is a cross-platform and open-source program (<http://www.mesquiteproject.org>) with most of the features of MacClade plus a broader array of methods, including independent contrasts.

**R** is a freely distributed program for statistical analysis and programming; individual users develop and contribute libraries that implement different methods (<http://www.r-project.org>). Several libraries are now available (*ape*, *ade4*, *geiger*, *PHYSIM*, *PHYLOGR*) that implement numerous phylogenetic comparative methods. *R* is a very powerful program that is being adopted by many researchers in ecology (although it is difficult to learn at first).

**COMPARE**, written by Emilia Martins, is a Web-based program that implements independent contrasts, phylogenetic linear models, and related methods (<http://www.indiana.edu/~martinsl/compare/>).

**Phylocom** is a freely distributed program (<http://www.phylodiversity.net/phylocom>) that is widely used for phylogenetic analysis of community structure and also conducts independent contrasts and analyses of phylogenetic signal.

**DISCRETE** and **Continuous**, both written by Mark Pagel and colleagues, implement several maximum-likelihood methods for the analysis of trait correlations, modes of trait evolution, and related methods. Both of these programs are now included in the BayesTraits program (<http://www.evolution.rdg.ac.uk/BayesTraits.html>).

### 3. PHYLOGENETIC SIGNAL: PATTERN AND SIGNIFICANCE

The fact that closely related species resemble each other—in ecological, morphological, behavioral, and other attributes—comes as no surprise to students of natural history. Evolution is generally a conservative process, and traits will usually change slowly, if at all, from one generation to the next. Adaptive radiations, in which species may diverge rapidly and take on novel adaptive traits and ecological lifestyles, are of interest precisely because they are unusual: at moments of ecological opportunity, following mass extinctions or the arrival of colonists on uninhabited islands, we see the potential for rapid evolutionary change. But most of the time, evolution is slow, and few changes accumulate, even over long periods of time. The lack of change is referred to as evolutionary stasis. The importance of understanding stasis in evolution has been highlighted by paleontologists, especially Steven Jay Gould, based on their study of the fossil record. When stasis, or at least a slow rate of change, plays out across the phylogeny, the result is that close relatives will be very similar.

Many terms have been used to describe this pattern of slow change: *phylogenetic inertia*, *phylogenetic constraint*, and *phylogenetic effects*. Often, these terms convey a sense that the phylogeny itself is the cause of ancestor–descendant resemblances. I find it useful to use the term *phylogenetic signal*, advocated in a recent essay by Simon Blomberg and Ted Garland, to emphasize that the similarity among relatives is a pattern and by itself does not reveal the underlying processes. An understanding of the causes of phylogenetic signal, and why it may vary in different groups and for different traits, draws on genetics, developmental biology, and ecology. We know that evolutionary change requires heritable, genetically based variation in a trait for selection to act on. Recent advances in the field of “evo-devo” are shedding light on how the process of development can influence the expression of genetic mutations, explaining why some traits vary more than others and why certain attributes may appear repeatedly in different lineages. On the other hand, even if ample genetic variation is available, natural selection may act to maintain traits in their current condition if an organism is well adapted to its current conditions. This process is known as stabilizing selection and may be pervasive in nature, although for a variety of technical reasons it can be quite hard to detect. The ability of plants and animals to migrate during episodes of climate change and track the environments to which they are well adapted may also be a process that reduces the rate of evolutionary change. There is no general consensus on the relative importance of these different factors that contribute to the phylogenetic signal in different traits, and it is very difficult to obtain all the relevant data in any particular case study.

In the context of ecological research, it can be useful to quantify the pattern of phylogenetic signal and compare observed patterns to those expected under alternative evolutionary models. The Brownian motion model, in particular, provides an important point of comparison because it is the foundation of many comparative methods. Although Brownian motion represents a random model of evolutionary change, it does generate a fairly high degree of phylogenetic signal, as sister taxa diverge gradually from their common ancestors. In contrast, null models in which trait values are randomly rearranged among the species in a study provide a baseline measure for the complete absence of phylogenetic signal. Two closely related measures, Pagel’s  $\lambda$  and Blomberg’s *K* statistic, are particularly useful, as they take on a value of 1 when patterns of trait similarity conform to expectations of Brownian motion and greater than or less than 1 when close relatives are more or less similar than expected, respectively. Another class of methods known as Mantel tests

is based on the correlation between the phylogenetic distances between species (the distance down the branches of the phylogeny to the common ancestor and back up to another species) and the ecological or phenotypic differences between them. These methods are useful for ecological characteristics such as niche overlap and co-occurrence where the degree of similarity or dissimilarity between species is quantified directly.

Phylogenetic information can play an important role in the prediction of ecological traits when there is strong phylogenetic signal. For example, in a recent study, Jérôme Chave and colleagues demonstrated that wood density tends to be very similar among closely related tree species. Wood density is important for carbon storage, a critical factor in the global carbon cycle, but it has only been measured on a small proportion of tree species in the tropics. Knowledge that close relatives have similar wood density will allow more accurate prediction of carbon storage in diverse tropical forests, even for species for which wood density has not been measured directly.

#### 4. PHYLOGENETICS AND COMMUNITY ECOLOGY

Phylogenetics is playing an increasingly important role in community ecology as a tool to gain insight into the processes that influence community structure. One of the earliest theoretical principles of ecology was the competitive exclusion theorem, formalized by Gause in the 1930s, which states that two species that utilize identical resources cannot coexist in a community. In the 1950s, this idea, together with the knowledge that closely related species are usually ecologically similar and therefore utilize similar resources, led to the prediction that species from the same genus should co-occur infrequently. This prediction was tested by calculating the average number of species per genus in isolated communities, such as islands, compared to the overall biota of the surrounding region. In the past 10 years, phylogenetic approaches to community ecology have been revitalized by the availability of highly resolved phylogenetic trees and new methods. In addition, developments in community assembly theory have emphasized an alternative view that co-occurring species may be more similar to each other than expected because similar traits may promote ecological success under particular environmental conditions. These two perspectives provide contrasting predictions regarding whether communities will be composed of more or less closely related species.

Three steps are required to quantify the phylogenetic structure of ecological communities and test hypotheses about whether this structure is significantly

different than may be expected. First, the degree of relatedness among co-occurring species needs to be quantified, based on the best available phylogeny. Cam Webb and others have introduced several related methods to accomplish this. The simplest approach is simply to calculate the average phylogenetic distance between all pairs of species within the community. Other approaches take into account species abundance or measure the distance between each species and its closest relatives in the community, as opposed to more distant relatives. The second step is to specify a broader pool of species from which a particular community has been assembled. This provides the source pool to construct hypothetical communities that serve as a point of comparison with observed patterns. Ideally, the spatial scale defining this pool is large enough so that it includes all of the species that could, in a reasonable span of time, arrive at the community of interest. However, in practice, it is very difficult to determine exactly what this scale should be, and researchers rely on a variety of practical solutions to address this problem. Finally, one needs to construct an appropriate null model by which random communities can be drawn from this regional pool to determine whether the observed communities diverge from random expectations. Simple null models include a random draw of species, where each species is equally likely to be chosen. More complex models can be constructed, in which the probability of a species being chosen is proportional to its frequency of occurrence in the landscape. The construction and analysis of these null models are continuing points of discussion and development in this field.

Many studies of phylogenetic community structure have appeared in recent years, and some generalizations are beginning to emerge. First, empirical and theoretical studies suggest an asymmetry in the interpretation of phylogenetic community data. It appears that clustering of close relatives within a community arises primarily from an ecological filtering process, in which similar species are favored as they share adaptations that are appropriate for the particular conditions. On the other hand, many different processes can lead to the opposite pattern in which communities are composed of more distant relatives than expected. These include competition, small-scale habitat heterogeneity, facilitative interactions among functionally disparate species, and even a filtering process when the traits that promote success have evolved independently in different clades. Theoretical studies also suggest that it is much harder to detect patterns in which coexisting species are distantly related, compared to the opposite pattern.

A second result is the realization that communities will not be structured either by filtering or by competition or by any other single process. Many processes

are likely at work, mediated by different sets of traits. For example, Jeannine Cavender-Bares and colleagues studied the composition of oak-dominated forests in Florida and found that local communities were generally composed of distantly related species. These species tended to share physiological traits affecting their water relations, with drought-adapted species occurring together on drier sites. Moreover, these hydraulic traits exhibited low phylogenetic signal, so similar species tended to be distantly related for these characteristics. On the other hand, co-occurring species displayed a high disparity of trait values related to acorn maturation and wood density. These traits exhibited a high degree of phylogenetic signal, but closely related species with similar values were distributed across different communities. Thus, it is critical to specify the traits that may be relevant to community assembly and examine their distribution on the phylogeny carefully before interpreting patterns of phylogenetic community structure in terms of particular underlying processes.

Finally, there is a fascinating pattern in plant communities of a shift from the co-occurrence of more distant relatives when studies focus on a narrow clade (e.g., oaks) to a pattern of clustering of close relatives in broader studies that encompass the full spectrum of flowering plants or all seed plants. A similar shift occurs moving from smaller to larger spatial scales. Both of these patterns are consistent with a stronger role for resource partitioning among closer relatives and at smaller spatial scales, whereas habitat filtering becomes more apparent at larger spatial and phylogenetic scales.

## 5. PROSPECTS FOR THE FUTURE

The potential role of phylogenetics in ecology was heralded by several articles and books published in the late 1980s to mid-1990s. In the relatively short interval since then, many methods have been introduced or improved, and growth in research has been rapid. The number of citations in the scientific literature under the keywords *phylogen* and *ecology* rose from 4 in 1990 to 87 in 1995, 130 in 2000, and 275 in 2006. An important engine of this growth has of course been the constantly expanding availability and improved resolution of phylogenies for diverse groups of taxa, accompanied by new methods, fast computers, and easy-to-use software. This chapter highlights two areas that

are most relevant to ecological research. Measures of phylogenetic diversity are also used as criteria to help prioritize taxa and habitats in conservation biology, and a wide variety of comparative methods are in use in evolutionary biology, including the study of adaptation, diversification, adaptive radiations, and related topics.

Several important areas of challenge and opportunity lie ahead. One is the improved resolution of branch lengths and node ages on phylogenies, which will be provided by including more species and more genes and improvements in fossil calibration. Time-calibrated phylogenies are opening the door to linkages between comparative methods and paleoecology and will facilitate investigation of a new generation of questions. A second area is the development of global databases for ecological traits. This will allow us to assess questions of phylogenetic signal and ecological trait correlations across the entire phylogeny of major clades and to understand how the assembly of local floras and faunas relate to global patterns of ecological diversity. Third, phylogenetic methods are providing new insights into ecology and biogeography of microbes, fungi, and other groups that are difficult to study directly in the field. These are but a few of the growth areas at the intersection of phylogeny and ecology—the most exciting advances will be those that at this point are not even anticipated.

## FURTHER READING

- Blomberg, S. P., and T. Garland, Jr. 2002. Tempo and mode in evolution: Phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology* 15: 899–910.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125: 1–15.
- Harvey, P. H., and M. Pagel. 1991. *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press.
- Maddison, W. P., and D. R. Maddison. 1992. *MacClade: Analysis of Phylogeny and Character Evolution*. Sunderland, MA: Sinauer Associates.
- Pagel, M. D. 1999. Inferring the historical patterns of biological evolution. *Nature* 401: 877–884.
- Webb, C. O., D. D. Ackerly, M. McPeck, and M. J. Donoghue. 2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics* 33: 475–505.

Copyright of Princeton Guide to Ecology is the property of Princeton University Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.