
Evolution of Behavior: Phylogeny and the Origin of Present-Day Diversity

TERRY J. ORD AND EMÍLIA P. MARTINS

Studying the evolution of behavior is a challenge. Behavior rarely leaves a fossil record, and selection experiments are generally feasible for only a limited number of short-lived insect and fish species. In his classic paper, Niko Tinbergen (1963) suggested two approaches to understanding behavioral evolution—studies of “survival value”, or function, and evolutionary history. Many behavioral ecologists interpret evolution through the first of these and emphasize the present-day function of behavior within populations. For example, we might find that the colorful plumage and courtship dances seen in male manakin birds attracts females, and that females in turn use these traits to assess and subsequently choose among several possible mates. Or perhaps we discover that Caribbean anole lizards producing strenuous physical displays during aggressive interactions do so to deter rivals from stealing territory. In each case, we might conclude that display behavior has evolved or, more specifically, is an adaptation for attracting mates or deterring rivals. But this is essentially a reiteration of function and tells us very little about how these behavioral traits actually arose and changed over time. In particular, it is difficult to decipher why a behavioral trait takes its present form or why there is so much diversity in functionally equivalent behavior across species. If a behavioral trait serves the same function, why does it vary in different species? Are these differences adaptive, or do they reflect some other evolutionary process? To

investigate evolution, we need to add Tinbergen’s second emphasis on the study of closely related species and consider the differences and similarities in behavior observed across species within the context of their evolutionary history. This is the foundation of the comparative method. A similar approach was adopted by Charles Darwin trying to explain the diversity of forms he observed in the natural world, which led him to develop his theory of evolution outlined in *The Origin of Species* (1859) and other classic works.

To understand fully the origins of behavior, a behavioral ecologist needs to consider that species are the product of their evolutionary past, in addition to present-day selective pressures (see chapter 2, this volume). Two ethologists noted for first implementing a rigorous scientific approach to the study of animal behavior, Konrad Lorenz and the aforementioned Niko Tinbergen, clearly advocated the explicit consideration of evolutionary history. Indeed, one of the earliest efforts to examine the evolution of behavior was Lorenz’s reconstruction of motor patterns and call types in waterfowl (figure 7.1). Without the aid of a computer, Lorenz meticulously grouped species together according to their similarities in behavior. He showed how some behavioral traits seemed to have evolved quite early in the group (i.e., traits shared by most species), whereas others were likely to be more recent innovations (traits unique to single or a subset of species). His main objective was to illustrate that

some behavioral traits could be as informative as morphology for distinguishing the evolutionary relationships between species. The fact that behavior is often retained from a common ancestor is an important observation and can help explain the form that many behavior patterns take in animals today.

The difficulty that these pioneering researchers and others faced in trying to use a comparative approach to study behavioral evolution was the need for large data sets of closely related species. Unlike other aspects of an animal's phenotype that can be quantified with more ease (e.g., morphology), collecting information on behavior is often labor intensive and time consuming, taking months or even years of work to compile for a single species. Ethologists were also limited by the lack of information on the nature of evolutionary relationships between species in many groups and robust methods for integrating this information into their investigations. With the progressive culmination of empirical research over the last half-century and the onset of the information age, it is now possible to compile large inventories of behavior and ecological information. The advent of molecular techniques heralded a proliferation of "tree building" in the 1980s and *phylogenies* (evolutionary trees that illustrate the relationships between species) are now available for many taxonomic groups. There is also a daunting array of computer programs, implementing a variety of different evolutionary models and handling a variety of different types of data. In short, behavioral ecologists today have an unprecedented opportunity to investigate the evolution of behavior within an explicit and increasingly sophisticated historical framework. In this chapter, we outline how the evolutionary history of organisms can be integrated into behavioral ecology research and the unique insights that can be gained in doing so.

CONCEPTS

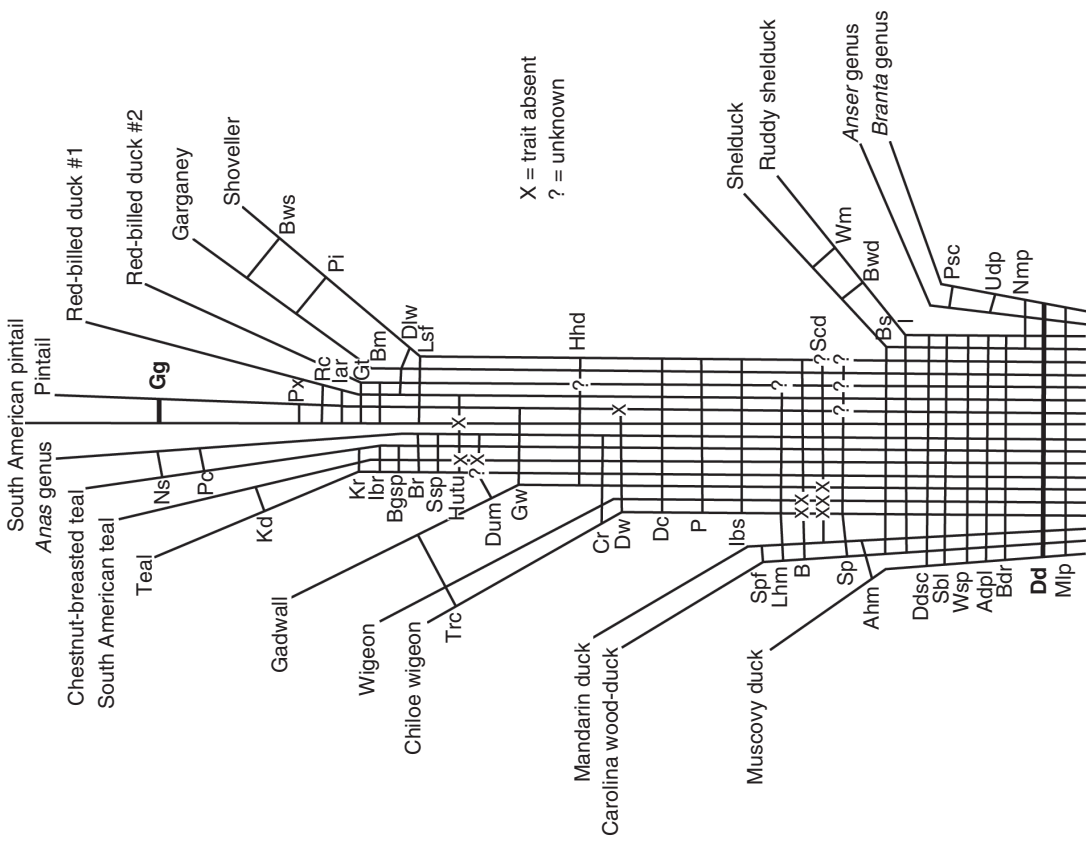
The study of the evolution of phenotypic traits that lack fossil records focuses on *the comparative method*, the principle of which is quite simple. We might observe that a behavioral trait in our favorite study organism seems uniquely suited or *adapted* to the environment that it currently lives in. If this is true, we might expect to see the same trait in other species living in similar environments, prompting

the obvious comparison of traits and habitats across species. This is the basic concept of the comparative method in its classical form (Darwin 1859). However, drawing inferences about adaptation from species comparisons is not as straightforward as it might first seem. What can we actually conclude from an observation that species exhibiting the same trait live in similar habitats? We have no way of knowing whether this pattern reflects independent evolution of the same trait in different species or whether the trait evolved once in a common ancestor and was subsequently retained in descendent species. In the latter scenario, a trait might have little or nothing to do with the current environment a species lives in. The modern revision to the comparative method incorporates information on the evolutionary relationships among species to distinguish examples of potential adaptation from other evolutionary processes.

Many aspects of the general appearance (morphology) and lifestyle (behavior and ecology) of closely related species are often shared because of common ancestry. For example, almost all Caribbean anole lizards produce extravagant head-bobbing displays in territorial defense. In the same way Lorenz concluded that behavioral traits shared by all waterfowl were ancestral, we can say that it is highly likely that the common ancestor to the anole lizards also defended territories and used head-bob displays. If we expand our comparison even further to the entire superfamily iguana (several thousand species across the globe), we find territoriality and visual displays are generally common to all, pushing the origin of these behaviors even further back in evolutionary time (Martins 1994; Ord et al. 2001). Even so, no two species are fully alike. Species might possess the same trait through common descent, but the trait will often differ to a small or large degree in its expression. For example, each lizard species has its own unique way of performing territorial displays. One species might use a series of rapid, shallow head bobs, whereas another produces protracted, high-amplitude head bobs (e.g., compare *Anolis sagrei* and *A. carolinensis* in figure 7.4). Although similarities in behavioral traits among closely related species can reflect the age of traits, the extent of modification or variation that exists in these traits can be equally informative.

How might we explain the diversity of iguanian territorial displays? By identifying ecological, morphological, and life history factors that vary among species in conjunction with the behavioral trait in

- Adpl: Anatinae duckling plumage
- Ahm: aiming head-movements as a mating prelude
- B: "burping"
- Bdr: bony drum on trachea
- Bgsp: black-gold-green speculum
- Bm: bill markings with spot and light-colored sides
- Br: bridling
- Bs: body-shaking courtship/demonstrative gesture
- Bwd: black-and-white duckling plumage
- Bws: blue wing secondaries
- Cr: chin raising
- Dc: descendo call of the female
- Dd: **display drinking**
- Ddsc: disyllabic duckling contact call
- Dlw: drake lacks whistle
- Dum: down-up movement
- Dw: drake whistle
- Gg: "geeeeegeeeee"
- Gt: graduated tail
- Gw: grunt-whistle
- Hhd: hind-head display
- Hutu: head-up-tail-up
- I: incitement by female
- Iar: incitement with anterior of body raised
- Ibr: isolated bridling not coupled to head-up-tail-up
- Ibs: introductory body shaking
- Kd: "koo-dick"
- Kr: "krick-whistle"
- Lhm: lateral head movement of the inciting female
- Lsf: lanceat-shaped shoulder feathers
- Mlp: monosyllabic "lost-piping"
- Nmp: neck dipping as a mating prelude
- Ns: nod swimming by female
- P: pumping as prelude to mating
- Pc: post-copulatory play with bridling and nod-swimming
- Pi: pumping as incitement
- Psc: polysyllabic gosling contact call
- Px: extension of the median tail feathers
- Rc: r-calls of the female in incitement and as contact call
- Sbi: sieve bill with horny lamellae
- Scd: social courtship
- Sp: sham preening, performed behind wings
- Spf: specific feather specializations serving sham-preening
- Ssp: speculum same in both sexes
- Trc: chin raising reminiscent of the triumph ceremony
- Udp: uniform duckling plumage
- Wm: black-and-white and red-brown wing marking
- Wsp: wing speculum



question, it is possible to generate hypotheses on the potential cause of trait variation in the context of phylogeny. A distinction needs first to be made between the *phylogenetic* comparative method and the approach employed in classical comparative psychology. Comparative psychologists study specific and very disparate model organisms (e.g., rat, pigeon, primate) to investigate the neurological or developmental pathways of behavior or to identify universal cognitive processes, such as those involved in learning. As is already becoming evident in this chapter, phylogenetic comparative biologists typically examine phenotypic variation in a group of closely related taxa, such as a genus or clade of birds, to investigate the evolutionary causes of phenotypic diversity (or similarity). Modern phylogenetic comparative methods were born in part from a statistical problem associated with analyzing this type of data, namely, that data points for closely related species are not independent of each other (Felsenstein 1985). Statistically, analyses such as regression or ANOVA make the assumption that for a given variable, the value of one datum is not related to the value of another datum (other than by factors explicitly described by the statistical model). This assumption is violated when data are compared across species because traits are often shared through common ancestry. Ignoring this constitutes a form of pseudoreplication and inflates the degrees of freedom in statistical tests. This leads to increased Type I errors in which a conclusion is drawn that an effect exists when in fact it does not (Martins 1996a).

FIGURE 7.1 The phylogenetic reconstruction of motor patterns, calls, and morphology in waterfowl adapted from Lorenz (1941). Vertical lines correspond to different species, whereas letters and horizontal lines are traits shared by species. For example, the “geeeeegeeeee” call, **Gg**, is exhibited only by two species of pintails, *Dafila spinicauda* and *Dafila acuta*. By grouping these two species with this common call type, Lorenz effectively reconstructs the geeeeegeeeee call as a recently derived behavioral trait. Compare this with display drinking, **Dd**, which is shared by all species of waterfowl and is subsequently placed at the base of the phylogeny. The modern-day interpretation of this assignment is that display drinking evolved early in the waterfowl group and has been subsequently retained in species today.

To alleviate this problem, Joseph Felsenstein (1985) suggested a novel way of removing phylogenetic nonindependence by transforming data using a procedure that has come to be known as *phylogenetic independent contrasts* (box 7.1). A variety of free, online computer programs are available that calculate these contrasts, and the technique remains the most popular phylogenetic comparative method in use today. Commentaries by Felsenstein (1985) and others (e.g., Ridley 1983; Cheverud & Dow 1985) also heralded a reevaluation of comparative analyses of interspecific data more generally (see Harvey & Pagel 1991; Martins 1996b). This prompted the development of a series of evolutionarily explicit techniques that not only incorporated phylogeny directly into analyses to account for species relationships but now go beyond simply “controlling” for statistical nonindependence of interspecific data to examining the process of evolution itself.

Although new methods are continuously being developed, most fall into three broad categories: reconstructing the ancestor states of traits, estimating the degree of *phylogenetic signal* in traits, and correlated trait evolution. We begin first by elaborating on what a phylogeny represents, how they are made, where behavioral ecologists can obtain them, and what are the underlying assumptions of using them to study the evolution of behavior.

Phylogeny: A Primer

A phylogeny is a hypothesis summarizing the evolutionary relationships between species. Phylogenies were initially developed through comparisons of organismal traits such as morphology, but most phylogenies today are based on mitochondrial and nuclear DNA. The details of how phylogenies are put together are complex and beyond the scope of this chapter (see Felsenstein 2004). Nonetheless, a behavioral ecologist wishing to use phylogenies to inform their research needs to have a basic understanding of the terms used. It should also be recognized that phylogenies are subject to change as new data comes to hand and the techniques used to create them are refined. This is an important point to remember because changes to the phylogeny can influence the outcome of a comparative analysis (e.g., see following section).

Phylogenies are depicted as trees (figure 7.2), with the names of extant species connected by a hierarchy of bifurcating *branches* summarizing

BOX 7.1 Comparative Methods

The majority of comparative approaches that behavioral ecologists will most likely use fall into two broad categories: techniques for reconstructing ancestor states and performing correlation analyses (e.g., for testing adaptive hypotheses). We provide a brief technical introduction to these approaches below.

Ancestral Reconstruction

Parsimony is an algorithmic procedure used to reconstruct phylogenies or to map ancestor states onto a phylogeny by favoring solutions requiring the least amount of evolutionary change in traits (figure 1a). In this sense, it tends to underestimate evolutionary changes and may not be as appropriate for behavioral evolution as for other aspects of the phenotype. Parsimony methods usually lack a statistical or probabilistic methodology; for example, assigning specific states following predefined rules without recognizing any possibility of uncertainty. Although it has come under fire for a lack of statistical rigor, and because it may not be a realistic view of how evolution occurs for many traits, parsimony approaches are often favored for their simple and straightforward computation.

Least Squares, Maximum Likelihood and Bayesian approaches are several newer methods that apply a probabilistic approach to find explicitly mathematical, evolutionary models that best fit the observed distribution of trait values across species (figure 1b). In the simplest case, these statistical methods yield results identical to those obtained using parsimony approaches. The difference being, however, that these methods also compare the probability

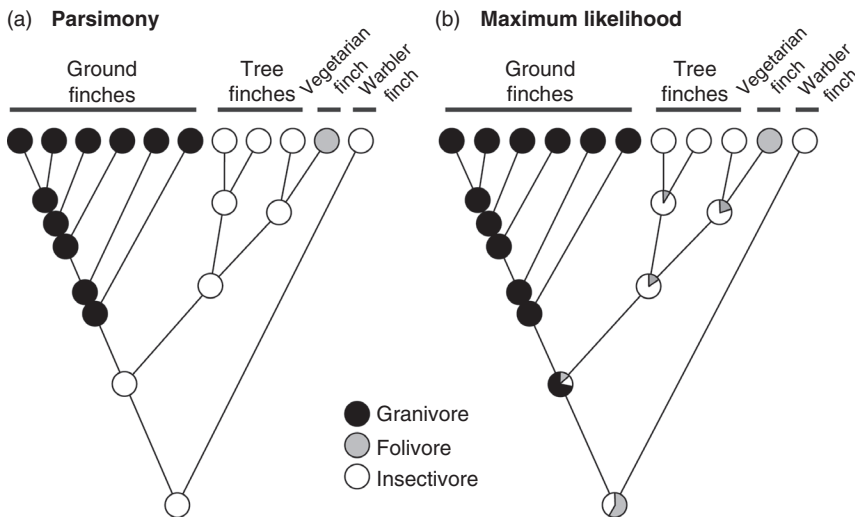


Figure 1 Alternative ancestral reconstructions of diet in Galapagos finches using parsimony and maximum likelihood methods. Adapted from Schluter et al. (1997). Pie charts reflect the relative support of dietary habits assigned to ancestors. The parsimony reconstruction suggests the root ancestor was an insectivore, whereas the maximum likelihood estimate implies it is more likely to have been a folivore.

(continued)

of various alternative scenarios of how evolution might occur—e.g., a model that describes trait variation as being explained by phylogenetic relationships among species (i.e., the null of Felsenstein’s approach), a model assuming no relation between trait variation and phylogeny (which is the same as not including phylogenetic information at all in analyses), and various models in between these two extremes (see Martins & Lamont 1998 for further discussion and examples in a behavioral context). The probabilistic framework has the added benefit of providing confidence intervals, emphasizing, for example, that the phenotypes of ancestors deeper in the tree are unlikely to be as well estimated as more recent ancestors.

Character Correlations and “Correcting” for Phylogeny

Not incorporating phylogeny into statistical analyses of comparative (interspecific) data can create a problem of pseudoreplication and statistical nonindependence. When estimating a general pattern across a large group of organisms, data measured from several closely related taxa should not be weighed as heavily as data from distantly related taxa. Because all taxa are related in one way or another, not incorporating phylogeny can lead to inflated rates of Type I error and the finding of significant effects when in fact none exist. By “correcting for phylogeny,” phylogenetic comparative methods (PCMs) correct the significance tests associated with parameter estimates for a fundamental statistical problem. In contrast, the absolute value of parameter estimates (e.g., means, correlation coefficients describing the relationship between traits) tends not to differ much, depending on the shape of the phylogeny or on whether a phylogeny is incorporated at all, unless the specific phylogeny being incorporated is extremely biased or unusual in some respect (see Martins & Housworth 2002; Rohlf 2006). The impact of incorporating phylogeny is primarily on the hypothesis test rather than on the value of the parameter being tested. Nevertheless, computer simulation studies (e.g., Martins et al. 2002) consistently confirm that PCMs nearly always perform better than not incorporating phylogeny, and even those who continue to debate exactly how and when PCMs should be applied (e.g., Rheindt et al. 2004) agree that the traits should be tested before a decision is made.

The phylogenetic independent contrasts approach is the most commonly used phylogenetic comparative method. Originally developed by Joseph Felsenstein (1985), this method corrects the basic statistical problems arising from phylogenetic relationships between species by transforming interspecific data into a set of differences or *contrasts* between immediate relatives, standardized by the amount of evolution expected along each branch of a known phylogeny (see figure II). If two sibling species express a trait value largely because of shared ancestry, then the difference score will be close to (or at) zero. Conversely, if the expression of a trait is unrelated to phylogeny, contrast values will be large. Although the computational algorithm requires estimating ancestral states along the way, these states are not intended to represent the phenotype of ancestors at phylogenetic nodes—ancestral states are better estimated using other techniques described above.

Felsenstein’s (1985) method assumes that phenotypic evolution is well described by a random walk or Brownian motion process, typically used by population geneticists to describe evolution under random genetic drift or fluctuating selection, and that the rate of evolution for the particular trait of interest along each branch of a phylogeny is known. A large family of related methods have been recently developed and can be interpreted as extensions of these assumptions, for example, allowing for a broader range of evolutionary scenarios (e.g., Pagel 1994; Martins & Hansen 1997; Butler & King 2004; Housworth et al. 2004), for not knowing the complete phylogeny or the exact rate or duration of evolution along each branch (e.g., Diaz-Uriarte & Garland 1998; Housworth & Martins 2001), and for testing whether phylogeny needs to be incorporated at all (e.g., Blomberg

(continued)

BOX 7.1 (cont.)

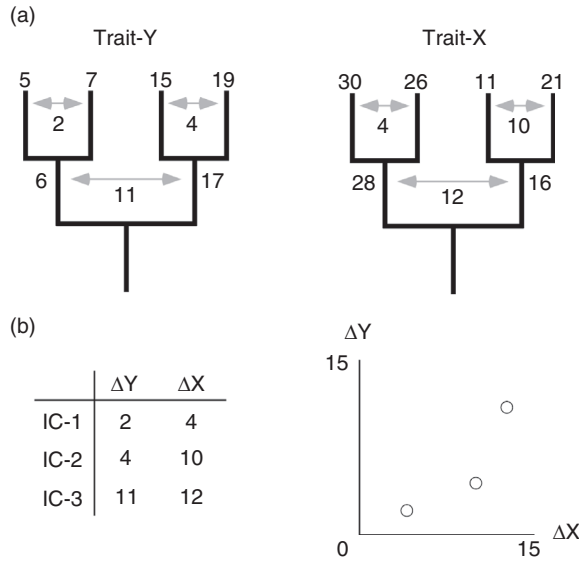


Figure II Stylized representation of the phylogenetic independent contrasts approach. (a) Difference scores or contrasts are calculated between species pairs and ancestor nodes on the phylogeny. (b) These contrasts are considered to be phylogenetically independent of each other and can be analyzed using standard statistical tests such as regression and ANOVA.

et al. 2003). These extensions provide additional accuracy and statistical power, clearly improving the fit and utility of Felsenstein’s original approach (e.g., Martins et al. 2002). New comparative methods also address a variety of novel questions. For example, the phylogenetic mixed model (Housworth et al. 2004) can be used to estimate and compare the degree of phylogenetic heritability in two traits. Hansen’s adaptation model (Hansen 1997; Butler & King 2004) estimates the degree to which a phenotype is adapted to different aspects of a complex selective regime. Others are using comparative approaches to understand how phenotypic evolution is related to species diversification (e.g., Ricklefs 2004). These new methods offer a powerful way to uncover evolutionary patterns that could not otherwise be observed.

the evolutionary relationships between species. Points at which branches split are called *nodes* and represent speciation events. A *polytomy* occurs when more than two lineages originate at a node. This can be indicative of rapid speciation (referred to as a *hard polytomy*) or when the precise relationship between lineages cannot be fully resolved (*soft polytomy*). In the latter instance,

this uncertainty will reduce the accuracy of comparative methods.

The length of phylogenetic branches often reflects the extent to which lineages have diverged. For example, Clade A shown in figure 7.2 has relatively short branch lengths, suggesting that the amount of divergence among these species is relatively low. Branch lengths are typically calculated from genetic

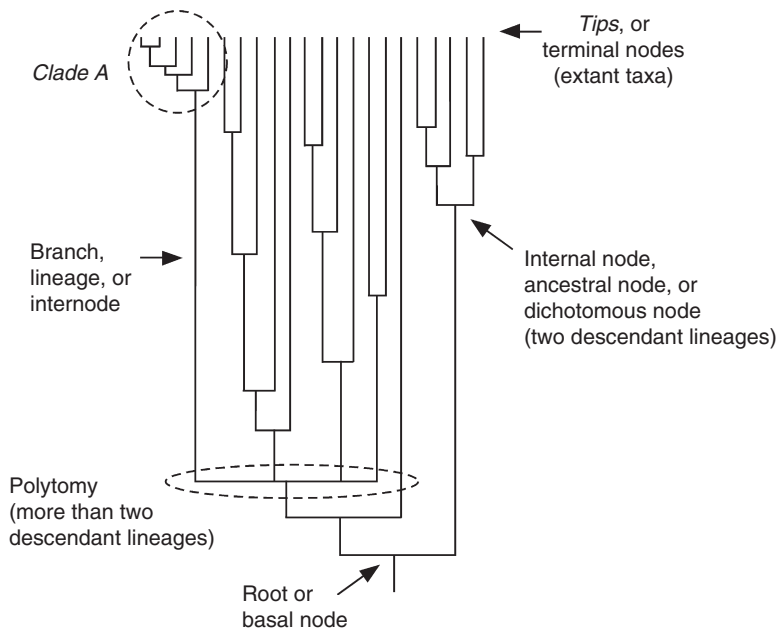


FIGURE 7.2 An example phylogeny illustrating common phylogenetic terms. Note that this and most phylogenies typically used in comparative analyses provide information only on evolutionary relationships among extant taxa and do not include extinct lineages.

sequence data, sometimes using *molecular clocks* or fossil evidence, to scale the rate of molecular divergence to units of real time (e.g., millions of years). Branch lengths are also sometimes reported in terms of number of morphological changes occurring along each branch. In comparative analyses, however, statistical methods require that branch lengths reflect the amount of evolutionary change expected for the behavioral trait being investigated. If we expect more evolutionary change along longer branches relative to shorter branches, then branch lengths in units of time may be a reasonable proxy for expected behavioral evolution. However, if we expect more change in the behavioral phenotype along particular branches of the phylogeny (perhaps because of a key innovation or adaptive radiation), change in the DNA sequences used to construct phylogenies may not correspond to rates of change occurring in the phenotypic traits under investigation, and the comparative method assumption will be violated.

There are a number of comparative methods, such as phylogenetic independent contrasts, that allow users to include information on branch lengths in analyses. In some cases, doing so assumes implicitly that trait evolution occurs via accumulated

incremental changes (Brownian motion or the so-called gradualism model of evolution; Darwin 1859). This continues to be the subject of considerable debate. Paleontological evidence (Eldredge & Gould 1972) and recent analyses measuring the nature of trait evolution (e.g., Pagel et al. 2006; Atkinson et al. 2008) suggest most evolutionary change occurs when new species are formed (i.e., at phylogenetic nodes) and that certain traits may evolve more quickly along certain branches than do others. Furthermore, estimates of branch lengths are subject to error (Felsenstein 2004). The alternative is to set all branch lengths equal to 1 (a punctuated model of evolutionary change), to try several possible sets of branch lengths generated by computer simulation or using additional information (e.g., Losos 1994; Martins 1996a), or to apply a more sophisticated comparative analysis that estimates relative branch lengths from the phenotypic data themselves (e.g., spatial autocorrelation; Gittleman & Kot 1990; phylogenetic generalized least squares method; Martins & Hansen 1997; Pagel 1997; phylogenetic mixed model; Housworth et al. 2004).

In most cases, phylogenies are obtained straight from the published literature. For a large comparative dataset, a single published phylogeny will

rarely cover all species of interest, requiring two or more trees to be merged together. Often this is done by extrapolating relationships between species with reference to those taxa that are common to all the sources being used. Branch length information may not be included in these composite trees because it is often difficult to know how to adequately standardize estimates between sources. With some gumption, a formal *supertree* can be constructed using genetic sequences compiled from databases such as GenBank (see Sanderson et al. 1998). Taking this approach requires some skill in phylogenetics.

Another common complication occurs when some species are not represented in any published phylogeny. One solution is to use taxonomic keys to position species. If there are large numbers in question, one remaining solution is to randomize the position of these taxa and conduct multiple analyses across all alternative phylogenies (e.g., Housworth & Martins 2001; Ord & Stuart-Fox 2006).

Ancestor Reconstructions

Fossil traces of behavior are rare, but it is still possible to infer the behavior of historical ancestors by reconstructing character states onto a phylogeny. Researchers use molecular phylogenies and a variety of computer programs to assign historical character states based on parsimony or probabilistic-based algorithms such as likelihood or Bayesian statistics (box 7.1). With these programs, it is possible to extrapolate how a particular ancestor might have lived from the behavior of descendent species that we can study today. A more powerful utility of ancestor reconstructions is the identification of broad macroevolutionary patterns. We can use phylogenetic trait reconstructions to discover whether a group of species exhibits a behavior because of common ancestry or whether it arose independently in different species through convergent evolution. The pattern of trait reconstructions on a phylogeny can in itself support or reject hypotheses on potential mechanisms or selective pressures believed to drive trait evolution (see case studies below). In particular, ancestor reconstructions are critical for testing hypotheses that assume a specific sequence of historical events.

Consider the hypothesis that preexisting biases in the sensory system of females promote the evolution and subsequent form of male traits used during courtship. This hypothesis explicitly predicts that the evolution of a trait occurs after the evolution

of a sensory bias. Phylogenetic studies are the only way to investigate the timing of these events. For example, males in several closely related Central American freshwater fish species are known as swordtails because they possess an elongated caudal fin or “sword.” This sword functions as an ornament for attracting females, with longer swords being preferred over shorter ones (Basolo 1990a). Surprisingly, females of several related species to the swordtails, the platyfish, also prefer males with swords even though males of their own species do not naturally develop them (Basolo 1990b). This prompted Alexandra Basolo (1990b) to test whether females had a preexisting bias toward males with long caudal fins, which leads to the clear prediction that female preference for swords should predate the evolution of the sword itself. Basolo used a phylogenetic reconstruction to pinpoint when the sword most likely arose in the group. She then showed through mate choice tests in females of several different swordtail and platyfish species that all exhibited some form of preference for swords, even in species in which conspecific males lacked swords. This information, combined with phylogenetic reconstructions, revealed that the sword evolved after the preference in females. The sword is also used by males in aggressive competition in several species, and additional phylogenetic studies confirm that both the sword (Moretz & Morris 2006) and female preference (Morris et al. 2007) evolved before males co-opted the sword as a signal for use in same-sex battles.

The swordtail example highlights several important considerations when attempting to reconstruct specific evolutionary events. Fresh scrutiny over whether the sequence of female preferences and the origin of male swords conforms to Basolo’s (1990b) original findings is brought with each revision to the phylogeny and development of alternative, arguably more evolutionarily realistic methods for assigning ancestor states (e.g., Meyer et al. 1994; Schuler et al. 1997; Marcus & McCune 1999). Indeed, the question has been revisited numerous times over the last 2 decades (e.g., Basolo 1991, 1996; Wiens & Morris 1996; Morris et al. 2007), but the sensory bias hypothesis has generally stood up to the test of time. As with any phylogenetic comparative method, ancestor reconstructions rely on a phylogeny and the underlying statistical framework of the method applied (box 7.1). When either is changed, so potentially does the outcome of ancestor reconstructions (Schuler et al. 1997; Losos 1999).

Moreover, the further back in time we attempt to reconstruct ancestor traits, the greater the level of uncertainty involved in estimates (Martins 1999). Results that are consistent across alternative phylogenetic hypotheses and using different methods for assigning states to ancestors will reflect particularly robust and prominent evolutionary trends.

Phylogenetic Signal

The variability observed in behavior within species sometimes leads to the naïve perception that all behavior must be highly evolutionary labile or, at the extreme, entirely plastic in origin. In other words, changes in behavior occur so rapidly—perhaps over a matter of generations or within a single generation (e.g., behavior that is acquired through experience, such as song learning)—that little or no trace of evolutionary history or phylogenetic signal will be seen in its expression today. The term *phylogenetic signal* specifically refers to the extent trait variation across extant species reflects underlying phylogenetic relationships. When traits exhibit no phylogenetic signal, it follows that reconstructing ancestor states or applying any phylogenetic comparative method is unwarranted.

Behavior is arguably the most dynamic aspect of an animal's phenotype. Of course, when more variation exists within rather than between species, a comparative approach is probably inappropriate. With this obvious exception, we can measure how much of the behavior observed across species is predicted by phylogenetic relationships and estimate statistically the propensity for evolutionary change in traits. Different aspects of behavior can exhibit a wide range of phylogenetic signal, from traits that are retained with little modification over long periods of evolutionary time, to those that exhibit large bursts of evolutionary change. Overall, behavior does tend to be more evolutionary labile than ecological, morphological, life history, and physiological traits (Blomberg et al. 2003), but behavior is increasingly being found to exhibit more phylogenetic signal than traditionally thought (Wimberger & de Queiroz 1996; Blomberg et al. 2003; Ord & Martins 2006). Ancestor reconstructions can provide an indication of how evolutionarily conservative a trait is by revealing whether a trait is homologous (shared through common ancestry). Methods that estimate phylogenetic signal provide a more accurate picture into the potential underlying evolutionary process. Traits that are closely tied

to phylogeny—have high estimates of phylogenetic signal or are homologous—could reflect a phenomenon that is sometimes referred to as *phylogenetic inertia*. This means little evolutionary change has occurred outside drift, perhaps because of low mutation rates, resulting in a trait that is retained more or less in the same form through descendent taxa. Phylogenetic inertia can also occur because of genetic, developmental, or physiological constraints limiting the amount of evolutionary change in a trait.

Alternatively, traits can appear relatively stable over long periods of evolutionary time through the entirely different phenomenon of stabilizing selection. For example, niche conservatism is the tendency for ecological traits in ancestors to carry over into descendent species, which in turn leads descendent species to settle in similar “preferred” habitat types as their ancestors. Behavioral traits that are ideally suited or adapted to a particular environment are subsequently maintained along diverging lineages (phylogenetic branches) via stabilizing selection. If the selection regime changes dramatically, either through the invasion of a novel habitat or a drastic change in the occupied habitat type (e.g., because of climatic change), either species will go extinct or adaptation will promote major evolutionary change. When two distantly related species occupy similar novel environments, this can lead to adaptive convergence. Depending on how recently and frequently these environmental changes have occurred, traits might show low or intermediate amounts of phylogenetic signal. Behavioral change that is plastic, such as behavior that is largely acquired through experience during an individual's lifetime, should intuitively have extremely little or no relationship with phylogeny. However, the extent to which a species is plastic in its behavior can be in itself an emergent trait of a species. The extent to which closely related species differ in their plasticity is an area of growing interest in evolutionary ecology and is an equally exciting avenue for future behavioral research (e.g., why are some species better learners than others?). Currently, we know little about whether behavioral plasticity as a species trait is or is not associated with phylogeny.

Correlated Trait Evolution

To establish that a specific selective pressure or constraint has led to behavioral change or stasis, we

need to confirm that variance in the putative causal factor (e.g., some aspect of the environment) corresponds with changes in behavior over evolutionary time. Although phylogenetic correlation tests are typically employed to infer adaptation, they can also determine whether morphology or some developmental factor has constrained evolutionary change in behavior. For example, the beak morphology of Darwin's finches on the Galapagos Islands is a classic example of ecological adaptation. Species that feed on large, hard seeds possess large, strong beaks, whereas those feeding on insects or smaller more easily opened seeds possess small short beaks (Grant & Grant 2008). However, beak morphology also influences the mechanics of song production, such that large beaks are less capable of performing rapid movement, and in turn influences the structure of vocalizations produced. A test of correlated trait evolution by Jeffrey Podos (2001) revealed a strong historical relationship between the evolution of large beaks and the production of songs with low syllable repetition and narrow bandwidths. Although song behavior is expected to be under directional selection from female mate choice or for greater efficiency in particular acoustic environments, the evolutionary diversification of beak morphology for exploiting different food sources has apparently constrained the range of song types that can ultimately evolve (Podos 2001).

This study of beak morphology in Darwin's finches employed the most common approach in studies examining coevolution, which is Felsenstein's phylogenetic independent contrasts (box 7.1). It assumes the null hypothesis that variation across closely related species is the product of phylogeny, meaning the consequence of stochastic processes such as genetic drift and mutation that culminates in phenotypic variation in the absence of selection. The contrasts resulting from such calculations are considered to be free of the "confounding" effect of phylogeny and potentially representative of adaptive change. There are a few important points to remember with this approach and the removal of all factors associated with phylogeny. First, this interpretation of Felsenstein contrasts incorrectly attempts to separate two inextricable aspects of the underlying process generating behavioral diversity: adaptation and nonadaptive processes. In the real world, we expect traits to evolve under a complex mix of the two processes working simultaneously. Second, our ability to detect adaptive phenomena is restricted to particular types of selection. Stabilizing

selection, for example, can result in adaptive evolution that tracks the phylogeny very closely (see previous section), as would constant directional selection that acts in the same direction across the phylogeny as a whole. Either of these forms of selection could be hidden if the causes of variation associated with phylogeny are not carefully considered and are subsequently removed rather than incorporated directly into a comparative analysis. Third, simplistic application of Felsenstein's method (e.g., as originally proposed) provides very little protection from possible errors in the phylogeny, branch lengths, and underlying evolutionary model (Martins et al. 2002). The calculation of individual contrasts is dependent on the position of the nodes and length of the branches being compared (box 7.1). Errors in either are subsequently magnified in the final estimate of the contrast. When branch length information is used, the scaling of contrasts assumes a Brownian motion model of evolution, which is meant to replicate gradual change in traits over evolutionary time. This is not necessarily how the traits being analyzed have in fact evolved. Most of the various extensions and additions of this method that have been developed more recently perform better statistically and are more robust to small errors (e.g., those implementing a generalized least squares approach; Martins & Hansen 1997; Pagel 1997). Fourth, Felsenstein's contrasts and many other comparative methods are explicitly designed to study the evolution of biological traits, and hence may not be appropriate for studies of species-level phenomena (e.g., degree of polygyny; Searcy et al. 1999) or environmental measures. For these, a technique that has been explicitly designed to measure the impact of environment on trait evolution may be preferred (e.g., Hansen 1997; Diniz-Filho et al. 2007). Fifth, animals evolve in a complex world and are subject to a plethora of concurrent evolutionary pressures. Behavioral phenotypes may not perfectly track a phylogeny or any single aspect of an environment because of competing selective pressures from other aspects of the physical environment (e.g., as envisioned in Hansen 1997, Butler & King 2004). For example, in lark buntings (Chaine & Lyon 2008), females prefer different aspects of male signals from year to year, choosing signals each year that best indicate a males' reproductive success in that year's particular combination of environmental features. Moreover, behavioral phenotypes often occur in *syndromes*, or groups of behavior types that are

linked to each other and to other aspects of the phenotype through genetic correlations, pleiotropy, or other underlying mechanisms (e.g., chapter 30). Thus, the evolution of behavioral phenotypes may not be clearly pegged to any single environmental feature.

Treating phylogeny as a statistical problem, rather than a direct contributing factor determining how species respond to selection, skews the philosophy behind the role adaptation plays in evolutionary diversification. Instead, behavioral ecologists may learn more from recent and more sophisticated comparative methods that adopt maximum likelihood as a way to integrate phylogenetic signal, phenotypic correlations, and complex environmental forces. For example, Thomas Hansen's (1997) method offers a means for estimating the relative impact of each of several possible selective pressures (e.g., humidity, temperature, degree of vegetation) on phenotypic evolution, assuming that the organisms are living in a complex selective regime in which all of these pressures play at least an occasional role. Hansen's method provides a way to compare the variance in a trait along lineages believed to have experienced stabilizing selection in a particular type of environment, to trait variances over the rest of the phylogeny (see also Butler & King 2004). Incorporating phylogeny directly into statistical tests rather than attempting to remove it from the data beforehand gives investigators a clearer picture of how behavioral traits evolve in the context of both adaptive and nonadaptive processes.

CASE STUDIES

Identifying Modes of Selection on Behavior

There are a number of different theoretical models that explain how female mate choice might drive or constrain the evolution of male display behavior. Empirical study can identify the presence and strength of female choice, as well as the traits exhibited by males on which females base their choices. However, such investigations within populations are unable to show the precise historical mechanism through which selection has acted. This is because different evolutionary processes can produce very similar traits, making it impossible to distinguish between different models of selection with

focused empirical research alone. One approach is to examine the macroevolutionary patterns of trait evolution using ancestor reconstructions. In many cases, different evolutionary processes will generate predictable patterns in how traits evolve along a phylogeny. By carefully considering the assumptions of different models, together with the natural history of a group, it is possible to use phylogenetic reconstructions to infer the predominant form of selection at work.

An early example of this approach is presented by Richard Prum's (1997) investigations of the evolution of male behavior and plumage in manakins. The group is comprised of 42 bird species and is known for its lek breeding system. Lekking males congregate at the same location and establish small display areas where they attempt to woo females through elaborate courtship displays. Although it is known that females choose among mates based on display attributes, the evolutionary mechanism behind this choosiness and how it drives differentiation of male traits across species is unclear. Prum (1997) examined two alternative sexual selection models and four hypotheses on how natural selection might contribute to display evolution in the group. By carefully considering the assumptions of each, he predicted several key macroevolutionary patterns that should be apparent once male displays are reconstructed onto a phylogeny. These predictions are summarized below and in table 7.1.

Fisherian or "Runaway" Male traits under Fisherian selection are arbitrary in the sense that they do not reflect condition or viability in males (although they may initially start off as indicator traits). Consider a population in which some males express long tail feathers and some females exhibit a preference for them. Males with long tail feathers will tend to have higher reproductive output because they are capable of acquiring matings from females both with and without the preference for long tails. On the other hand, males without the trait can mate only with females exerting no preference. Assuming that both the expression of the trait and its preference are genetically determined, they will become linked over successive generations. This mutual reinforcement will accelerate the proliferation of both the trait and preference throughout the population in a "runaway" process (Fisher 1930) and may ultimately lead to the exaggeration of the male trait beyond its initial form (chapter 24).

TABLE 7.1 Predicted macroevolutionary patterns in male behavior and plumage under different forms of selection

Mechanisms of Selection	Historically Nested Distribution	Skewed Distribution	Likelihood of Convergence	Diversity of Repertoire	Correlates of Divergence
Sexual selection					
Fisherian	Yes	Toward tips	Low	High	
Quality indicator	No	Toward base	High	Low	
Natural selection					
Mate choice efficiency			High	Potentially high	Environment (e.g., predation, search costs)
Species recognition	No		Low	Low	Sympatry
Sensory bias	Yes		High	Low High	
Narrow					
Broad					
Sensory drive	Yes		High	Potentially high	Environment (e.g., habitat structure, predation pressure)

Source: Prum (1997).

Consequently, these traits are expected to evolve freely in direction and elaboration. Prum argued that this should result in the evolution of multiple traits and diverse display repertoires across species, with few instances of convergence. The distribution of traits reconstructed on the phylogeny would also tend to be historically nested, in which traits are retained within clades and display repertoires increase as additional traits evolve (figure 7.3a). Fisherian selection also predicts rapid and continued diversification of male traits. Hence Prum argued that Fisherian selection will tend to skew trait evolution toward the tips of the phylogeny (figure 7.3d).

Quality Indicators Traits evolving under this form of selection are tied to male condition or viability. The number of possible traits that will reliably convey this type of information is limited, and Prum proposed that this should promote frequent convergence and little diversity in display repertoires across species. The potential for evolutionary change is constrained because quality indicators must provide honest information on condition and are therefore not as free to vary as Fisherian traits. Indicator traits should instead be retained with little modification in descendent taxa and, because female mate choice

is believed to be ancestral in manakins, such traits should have evolved early in the group and this will tend to skew trait reconstructions toward the base of the phylogeny (figure 7.3c). Alternatively, novel traits might arise that convey more accurate information on male quality. These will replace earlier forms because indicator traits are costly for males to maintain and redundant traits will be selected against. Prum suggested that this “trait switching” will lead to a nonhistorically nested distribution and also reduce the diversity of display repertoires observed across species (figure 7.3b). A repertoire of indicator traits would evolve only if each provides unique information on different aspects of condition. However, because there is a limited pool of potential indicator traits, the evolution of unique traits will progressively decrease with increasing repertoire size and this will also skew trait evolution toward the base of the phylogeny.

Mate Choice Efficiency The longer it takes for females to find and choose among males, the greater the cost imposed by natural selection in the form of energetic expenditure and predation pressure. The strength of these selective pressures is dependent on the ecology and environment of a species. Prum thus predicted convergences in display traits among

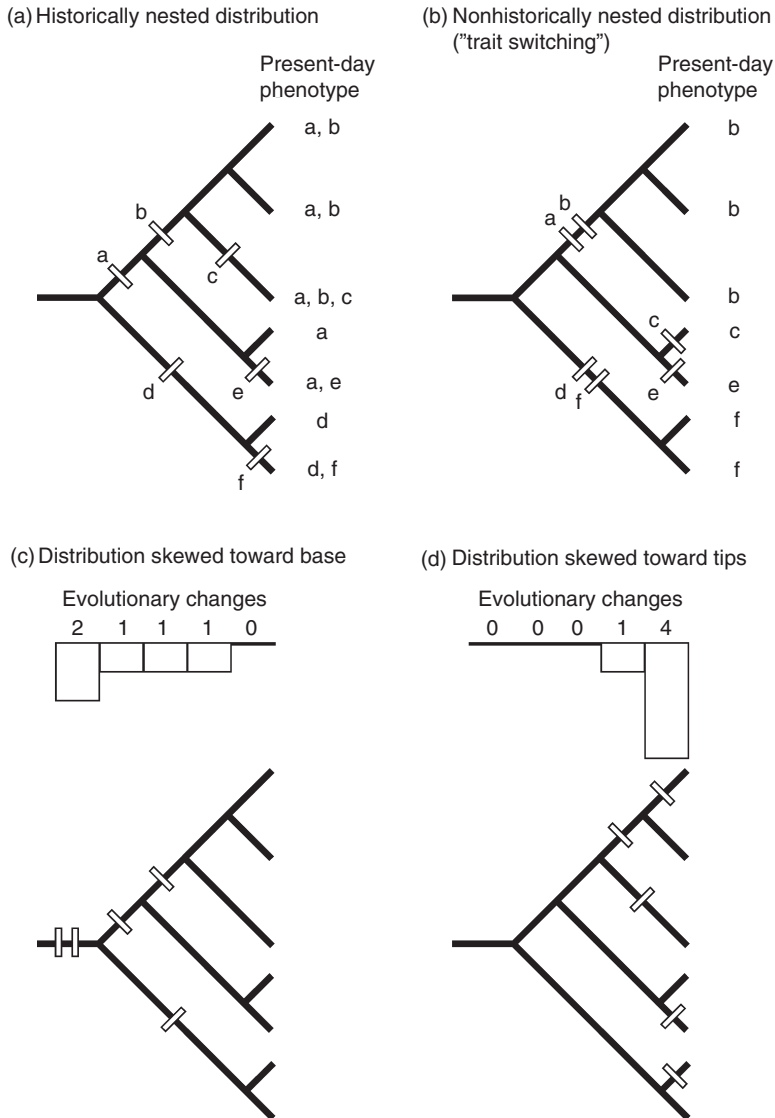


FIGURE 7.3 Hypothetical phylogenetic distributions of trait evolution modified from Prum (1997). (a) Traits exhibiting a historically nested distribution in which traits are retained within a clade in addition to the evolution of subsequent novel traits. This distribution should also lead to diverse repertoires and is predicted under Fisherian and sensory bias mechanisms of selection. (b) Traits exhibiting a nonhistorically nested distribution in which the evolution of novel traits replace those that arose earlier in a lineage's history. This distribution should lead to simple repertoires and is predicted when male traits function as quality indicators and/or in species recognition. (c) Trait evolution skewed toward the base of a phylogeny. This is predicted when traits are subject to some form of evolutionary constraint, such as traits that function as quality indicators. (d) Trait evolution skewed toward the tips of the phylogeny. This is predicted for traits that have not been constrained and have subsequently experienced rapid diversification, such as male traits subject to Fisherian or runaway selection.

species with similar ecologies and/or occupying similar habitats. Conversely, changes in the environment should promote the evolution of unique display traits. Diverse repertoires can also evolve if multiple traits contribute to mate choice efficiency or female survival.

Species Recognition Females are expected to be under considerable selection for discriminating among heterospecific and conspecific males. This would predict trait differentiation when species are sympatric and select against shared, homologous traits, which would confuse mate recognition. Prum suggests that this will constrain the evolution of diverse display repertoires, reduce the historically nested pattern in trait reconstructions (figure 7.3b), and limit display convergence.

Sensory Bias and Sensory Drive In these two related models, trait evolution is arbitrary and exploits a preexisting bias in the sensory system of females. The evolution of male traits that successfully tap into narrow biases are likely to be infrequent because the probability of such traits arising in the first place is low. Prum notes that if they do evolve, evolutionary change will be limited because trait properties are constrained by the specificity of the bias. On the other hand, a broad sensory bias could lead to the evolution of a number of different novel traits, accumulating into a diverse, historically nested display repertoire (figure 7.3a). The sensory drive model is the combined effect of sensory biases, habitat properties dictating signal efficiency and pressures from predators or parasites. Like the mate choice efficiency model, male traits are expected to diverge or converge depending on the differences or similarities in the ecology and environment of different lineages.

With these predictions in hand, Prum (1997) reconstructed male display repertoires and plumage traits onto the manakin phylogeny using parsimony. He also replicated the analysis for the same types of traits in a largely monogamous group of tyrant flycatchers, in which both sexes typically contribute to the care of offspring and in which the criteria used by females to choose between mates is potentially quite different. By doing this, Prum could compare the macroevolutionary patterns in two very disparate groups: a polygynous lek breeding system and a monogamous, biparental care breeding system.

Phylogenetic reconstructions for the manakins revealed that male traits are historically nested and generally skewed toward the tips of the phylogeny. Convergences in behavior or plumage characteristics have been rare, and display repertoires are extremely diverse across the group as a whole. Based on this evidence, Prum (1997) argued that Fisherian selection, or potentially a broad sensory bias, has been the predominant form of selection driving the evolution of extravagant male traits in manakins. By comparison, trait evolution in male flycatchers was not historically nested and display repertoires are quite simple. This suggests that male display traits in flycatchers have been selected as quality indicators, which is consistent with the monogamous breeding system of the group (Prum 1997).

Of course, the results of this analysis are dependent on the phylogeny and reconstruction method employed. The location at which specific traits are reconstructed onto the phylogeny will vary with changes made to the phylogeny or the algorithm used to assign ancestor states (e.g., compare the analyses of Basolo 1990b, 1991, Meyer et al. 1994, and Schluter et al. 1997 for the swordtail example mentioned earlier). However, because Prum (1997) examines major macroevolutionary patterns for dozens of traits collectively (e.g., 44 display behaviors and 44 plumage traits in the case of manakins), it is unlikely that minor changes in phylogenies or methods would lead to any dramatic changes in the overall reconstructed distributions observed. Other issues arise with the way traits are initially defined, which will also affect the type of reconstructed patterns observed on a phylogeny. For example, is a feather ornament on the same part of the body that varies in color among species a homologous single trait (one of the same origin) or several different traits? There is also an implicit assumption being made that the mode of selection acting on traits is largely consistent across species. It is possible that different selective pressures have acted on the same homologous trait in different lineages, leading to different rates and types of change. This is unlikely to leave any consistent patterns across a phylogeny (which in itself would be informative).

Identifying by empirical means the selective mechanism promoting the evolution of behavior can be extremely difficult. Indeed, Prum (1997) presents a compelling argument against extrapolating evolutionary processes from the adaptive function of a behavior as observed within a population

today (Gould & Vrba 1982; Baum & Larson 1991). Although empirical evidence might offer support for the hypothesis that a trait conveys honest information on male condition, it may still have originated under Fisherian selection and subsequently become linked to male quality over evolutionary time. As the manakin/flycatcher example demonstrates, selection can leave telltale signs in macroevolutionary patterns and reconstructing ancestor states is a powerful way to identify the form of these broad patterns.

Multifaceted Behavior

Individual traits do not evolve in isolation from the rest of the phenotype. Traits are often linked to each other developmentally, physiologically, or genetically. Even if traits are hypothetically free to vary independently of each other, selection acts collectively on an animal's overall phenotype and not literally on a single trait in isolation. The intrinsic factors influencing the evolution of behavior are likely to be as intricate as the specific selective pressures acting on behavior externally. An added challenge faced by researchers attempting to study a suite of traits is whether each feature being measured actually represents a separate trait or is an element of the same trait.

We can clarify the suitability of measures taken through experimental study of function. But a researcher still faces the problem of determining whether one key trait is the primary target of selection while others covary via some underlying genetic or developmental dependency, or whether each trait is under direct selection. Selection experiments in the laboratory are an excellent means to separate out intrinsic correlations and discover the focus of selection pressures (e.g., Chenoweth & Blows 2003; Mackay et al. 2005), but even in systems in which these experiments are possible (there are few in behavioral ecology), we are studying microevolutionary processes (i.e., within populations) that might not translate fully to the macroevolutionary scale (across species).

In a recent study, we investigated the evolution of a multifaceted behavior in the form of visual displays performed by West Indian anole lizards (Ord & Martins 2006). Males compete for territories using rapid up/down movements of the head, termed *head-bob displays*, and the extension of a throat fan known as a *dewlap*. The way lizards perform these displays varies across species extensively

in the number, duration, and type of bobs/dewlaps incorporated (figure 7.4). This diversity suggests anole displays constitute multiple traits that are at least semi-independent of each other, with each trait potentially targeted by different forms of selection from the social and physical environment or by other factors (Ord & Martins 2006). For example, variance in the intensity of competition over territories and the need to convey species identity are both suggested to explain signal diversity across species (Jenssen 1977, 1978; Ord et al. 2001). Anoles also occupy a variety of different habitat types, and signal theory predicts animals will evolve signal forms that enhance transmission efficiency in the environment they typically communicate (Endler 1992). Each of these selective processes is unlikely to influence all signal traits equally. Genetic correlations will further complicate evolutionary outcomes, and differentiation between species is also possible in the absence of selection through mutation and genetic drift.

To explore these factors, we scored 15 display features for 53 taxa using published plots that summarize the structure of displays (e.g., figure 7.4). To highlight the presence of trait correlations, we entered display data into a principal component analysis that incorporated basic phylogenetic effects using Felsenstein's independent contrasts. This revealed several strong associations among display traits that enable us to identify key display features that summarize the diversity observed across species (and were unlikely to measure the same trait), as well as trait complexes that have potentially evolved in concert (table 7.2). To describe the relative evolutionary lability of display traits, we used two different maximum likelihood approaches to estimate how freely each trait has undergone potential adaptive change and the amount of trait variation (adaptive or otherwise) inherited directly from evolutionary ancestors (Lynch 1991; Martins & Hansen 1997; Housworth et al. 2004). We found two traits relating to how dewlap displays are performed to be tightly bound to phylogeny, indicating that variation in these traits is explained to a large degree by the evolutionary relationships between species. Change in these traits (table 7.2) has been gradual, reflecting either random genetic differentiation or some factor tracking phylogeny (e.g., stabilizing selection or physiological constraints). The majority of display traits, however, showed very little phylogenetic signal, suggesting extremely rapid

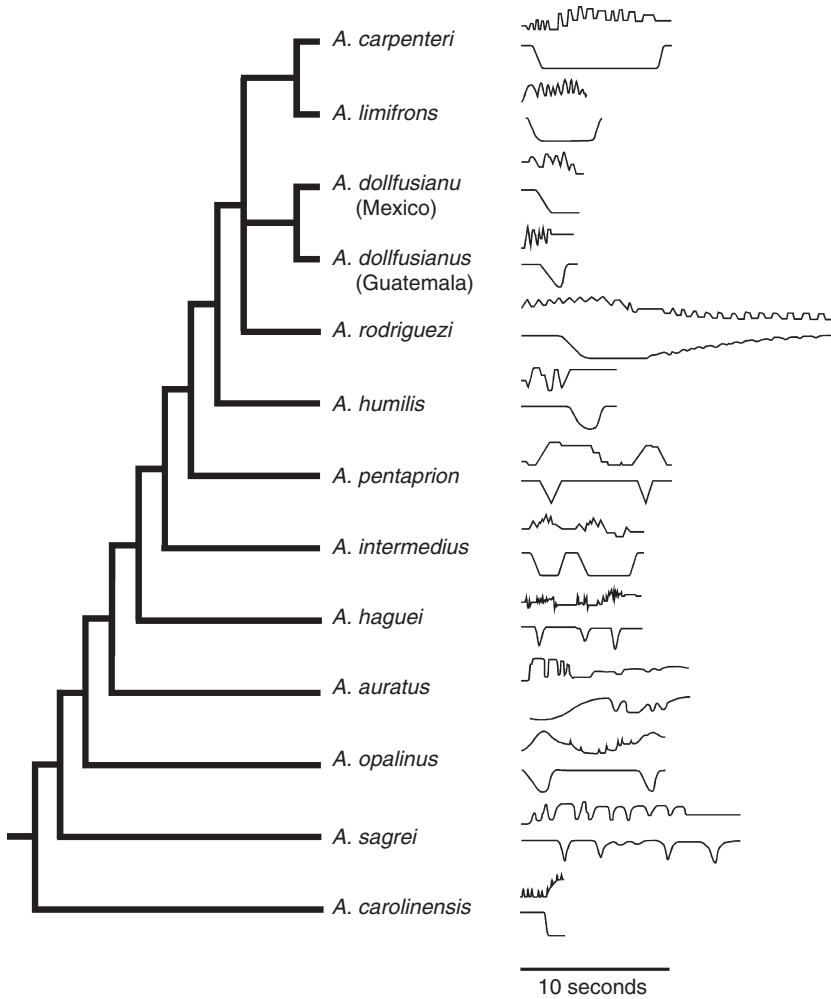


FIGURE 7.4 A subset of *Anolis* species illustrating the diversity in territorial displays. Plots represent the up/down movement of the head-bob (top line) and the extension/retraction of the dewlap (bottom line) over time. See Ord and Martins (2006) for sources used to draw display plots and phylogeny.

rates of change. We also examined several ecological traits and found that the type of habitat species occupied showed almost no phylogenetic signal (Ord & Martins 2006), confirming the absence of niche conservatism in these animals (Losos et al. 2003). Other ecological traits reflecting competitive intensity and sympatry were moderately tied to phylogeny (Ord & Martins 2006).

To explore more directly the presence of adaptive change in displays, we used several different phylogenetic regression models and found consistent evidence that competition among males, species recognition, and habitat use have all apparently

promoted adaptation in different aspects of display behavior. Anoles experiencing intense competition seem to have evolved short, complex displays, which parallels what has been suggested for the evolution of bird song (e.g., Read & Weary 1992; MacDougall-Shackleton 1997). Sympatric anoles possess displays that are complex in other aspects, supporting the view that territorial displays also function in species recognition. Finally, despite some dewlap behavior being evolutionarily conservative, change in the design of this aspect of the display has nevertheless still occurred following the invasion of new habitats.

TABLE 7.2 Trait complexes and estimates of phylogenetic signal for territorial displays in anole lizards

	Trait Complexes					Phylogenetic Signal	
	1	2	3	4	5	α	h^2
Head-bob duration	+					12.38	0.25
Dewlap duration	+					15.34	0.09
Average head-bob duration	+						
Average dewlap duration	+						
Average head-bob pause duration	+	-					
Average dewlap pause duration		-					
Dewlap latency		-				6.19	0.90
% overlap		+					
Head-bob number			+			12.99	0.05
Dewlap number			+			15.33	0.02
Dewlap amplitude variation			-				
Head-bob amplitude variation				+			
Dewlap/head-bob ratio				+	-		
Head-bob uniformity				-		14.85	0.19
Dewlap uniformity					+	8.91	0.60

Note: Principal component analysis on Felsenstein's independent contrasts revealed five primary trait complexes, suggesting these traits might have evolved in concert. Signs indicate the direction of evolutionary relation between traits (positive or negative). For example, increases in the variation of head-bob amplitudes included in displays correspond to decreases in the temporal uniformity of head bobs, a trend that has been documented in other lizard groups (Martins 1993). Two estimates of phylogenetic signal are also presented for key display traits in each trait complex: α represents how free a trait is to vary along the phylogeny, with larger values indicating greater levels of change; h^2 is the extent traits are phylogenetically heritable, with high values indicating a strong correlation to phylogeny. See Ord and Martins (2006) for details on display traits scored.

We can therefore explain the evolution of diversity in anole territorial displays on two fronts. First, displays are made up of several traits, each differing in the propensity for evolutionary change. Second, different selective pressures have targeted different traits, producing changes in behavior that have either been extremely rapid or incremental over long periods of evolutionary time. By adopting several different phylogenetic approaches, we were able to show that the evolution of complex behavior is the product of an equally complex interaction between factors associated with phylogeny, correlations between traits, and an elaborate selective regime.

Thinking Phylogenetically

Even when large data sets of behavioral information are not available to a behavioral ecologist,

phylogeny can still be highly informative in explaining the present-day function of behavior. Indeed, being explicitly aware that behavior carries the imprint of past selective pressures, as well as other factors associated with phylogeny, can be the only way to interpret why animals behave the way they do. For example, antipredator behavior is often assumed to be an adaptive response to predators in the current environment. However, Sih et al. (2000) observed this was not the case for the streamside salamander. Compared to other species, the high activity of larval streamside salamanders make them conspicuous targets for sunfish, their current predator, and are generally too quick to emerge from refuges when hiding from these fish predators, which increases the probability that predators are still present on emergence. Sih et al. (2000) provided an explanation for this maladaptive behavior by invoking phylogeny.

Those salamander species exhibiting effective tactics for predator avoidance (low activity, conservative emergence times from refuges) are also species known to have a long history with fish predators. The streamside salamander, however, is believed to have diverged from a recent ancestor species that was unlikely to have been exposed to fish. Sih et al. (2000) provides evidence to support the view that this predator-naive evolutionary history has subsequently limited the capacity of streamside salamanders to respond appropriately to the predatory environment they currently find themselves in.

Behavior can also be inherited from ancestors and still expressed despite serving no current adaptive function. Cowbirds and cuckoos lay eggs in the nests of other bird species to avoid the considerable investment involved in raising young. Many target species of this parasitism have evolved ways to identify and eject these foreign eggs from their nests. Rothstein (2001) discovered that the loggerhead shrike in California exhibits this same egg rejection behavior when he placed fake eggs into their nests. The surprise comes when we realize that the loggerhead shrike is not currently parasitized by other birds, meaning that this rejection behavior has no function. Rothstein was aware of the phylogeny of loggerheads and able to explain the existence of this behavior because he knew that loggerheads are nested within a large, Old World clade, all the species of which are believed to be parasitized and express egg ejection behavior. In other words, loggerheads express superfluous ejection behavior because they inherited it from evolutionary ancestors, not because it serves any adaptive function today.

Neither of these studies applies any formal phylogenetic comparative method. Nonetheless, the authors explicitly incorporate phylogeny directly into the interpretation of their experimental findings. Thinking phylogenetically, or considering carefully the phylogeny of a species being investigated, forces an investigator to acknowledge that the current environment might not necessarily explain the origin or even the precise function of the behavior being examined.

FUTURE DIRECTIONS

Behavioral phenotypes are intrinsically complex. As our ability improves to detect and study the

mechanisms underlying the production and ontogeny of behavior and the environmental factors that guide behavioral evolution, it becomes increasingly difficult to make sense of an overwhelming diversity of environmental, physiological, and social factors that influence behavior. The comparative method offers a powerful tool for honing in on specific selective factors and behavioral aspects that have evolutionary consequence. For example, studies comparing behavioral syndromes in different fish populations (e.g., Bell 2005; Moretz et al. 2007) have found intriguing evidence that selection acts simultaneously on suites of behavioral traits at the population level. Phylogenetic comparative studies further exploring the habitat features that influence the evolution of such correlated suites of behavior at higher taxonomic levels promise special insight into which aspects of a complex selective regime merit special attention. Similarly, comparisons of gene expression in finch species with different levels of social behavior (Goodson et al. 2005), of endocrine response in birds with different mating systems and levels of parental care (Hirschenhauser et al. 2003), and of the genetic patterning underlying development of insect social behavior (e.g., Toth & Robinson 2007) have helped narrow down the number of genes, hormones, and receptors involved in the evolution of complex behavior.

Phylogenetic approaches are increasingly being used in sexual selection research. The study of sensory biases in mating behavior often necessitates it (e.g., Basolo 1990b). Others interested in how sexual selection might promote speciation have begun to apply phylogenetic comparative methods to estimate rates of species diversification in clades that have experienced strong sexual selection relative to those that have not (e.g., Gage et al. 2002). As Prum's (1997) study illustrates, the form that sexual selection takes is an important question and should be revisited. In particular, the recent proposal that competing models of sexual selection lie at different ends of the same continuum (Kokko et al. 2002) might be examined within the context of phylogeny and potential ecological factors that contribute to divergences in the type of selection experienced by species. Something similar could be attempted for alternative game theory models and the evolution of animal contests (chapter 15).

Confirmation through some form of phylogenetic study provides persuasive support for a behavioral trait considered to be an adaptation.

Examples might include testing the correlated evolution of mating systems, antipredator strategies, kin selection, or altruistic behavior between unrelated individuals, with historical changes in social, predatory/parasitic, or physical environments. Explaining the origins of sexual dimorphism can also be examined phylogenetically. For example, comparative studies can estimate the extent traits expressed by the sexes are historically correlated or free to vary independently, and whether the selective regimes acting on traits differ between the sexes (e.g., Stuart-Fox & Ord 2004; Ord & Stuart-Fox 2006).

We still know little about the rate and type of evolutionary change (punctuated equilibrium or gradualism) experienced by behavioral traits and whether it matches similar patterns in other phenotypic traits like morphology (Blomberg et al. 2003). Furthermore, what is the consequence of phenotypic plasticity and individual variation for macroevolutionary change? There is some debate over whether the flexibility of behavior that allows the invasion of new habitats or shifts in the use of existing habitats promotes or constrains evolution (chapter 6). With careful thought, these two alternative outcomes could be teased apart using a phylogenetic analysis estimating the extent to which traits are variable within, relative to across, species and the consequences of historical shifts in habitat use on rates of change. In addition, the extent to which species are plastic in their behavior because it is a selected trait and/or inherited from phylogenetic ancestors is unknown, yet is at the heart of understanding how plasticity influences macroevolutionary change in behavior and other phenotypic traits across species. One approach would be to quantify within species the extent to which a behavior is plastic to certain changes in the social or physical environment and repeat this for a number of closely related species to estimate the phylogenetic signal of plasticity as a species trait and what ecological/historical factors might correlate with its presence in some species and not others.

More generally, behavioral ecologists need to think more explicitly about evolutionary history and the role phylogeny plays in shaping present-day behavior. The emphasis on current utility is still prevalent in behavioral ecology. Although it is unlikely anybody would disagree with the notion that the evolutionary history of a species is important, it tends to be a post hoc discussion point rather

than an intrinsic element in behavioral research. By emphasizing phylogeny during empirical investigation, researchers may become aware of better experimental designs, clearer explanations to puzzling behavioral phenomenon, or perhaps even more fundamental questions to address.

SUGGESTIONS FOR FURTHER READING

The arguments presented by Gould and Lewontin (1979) are still pertinent today. Their paper, and several responding papers that followed (see chapter 2), should be required reading for all students of behavioral ecology. The body of literature on phylogenetic comparative methodology is vast. We suggest papers by Pagel (1999), Martins (2000), Nunn and Barton (2001), and Omland et al. (2008) as starting points. There are also several books and edited volumes available. Most are somewhat dated in the specific techniques described, but they still provide a good introduction to the concepts of comparative biology and historical perspectives for studying evolution. We recommend *The Comparative Method in Evolutionary Biology* (Harvey & Pagel 1991), and *Phylogenies and the Comparative Method in Animal Behavior* (Martins, 1996b). Other useful texts include *Phylogeny, Ecology, and Behavior* (Brooks & McLennan 1991), and *New Uses for New Phylogenies* (Harvey et al. 1996). Finally, *Inferring Phylogenies* (Felsenstein 2004) provides the most recent text and comprehensive review of phylogenetic techniques. It largely focuses on approaches used to construct phylogenetic trees, but the techniques described (parsimony, likelihood, Bayesian approaches) are the same as those employed in reconstructing ancestor character states and many phylogenetic correlation tests.

- Brooks DR & McLennan DA (1991) *Phylogeny, Ecology, and Behaviour*. Univ Chicago Press, Chicago, IL.
- Felsenstein J (2004) *Inferring Phylogenies*. Sinauer Associates, Sunderland, MA.
- Gould SJ & Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond B* 205: 581–598.
- Harvey PH, Brown AJL, Smith JM, & Nee S (eds) (1996) *New Uses for New Phylogenies*. Oxford Univ Press, Oxford.

- Harvey PH & Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford Univ Press, New York.
- Martins EP (1996b) *Phylogenies and the Comparative Method in Animal Behaviour*. Oxford Univ Press, New York.
- Martins EP (2000) Adaptation and the comparative method. *Trends Ecol Evol* 15: 296–299.
- Nunn CL & Barton RA (2001) Comparative methods for studying primate adaptation and allometry. *Evol Anthropol* 10: 81–98.
- Omland KE, Cook LG, & Crisp MD (2008) Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. *BioEssays* 30: 1–14.
- Pagel M (1999) Inferring the historical patterns of biological evolution. *Nature* 401: 877–888.