• Biodiversity on Federal Lands • Chinese Invasives

BioScience

Organisms from Molecules to the Environment

April 2008 American Institute of Biological Sciences Vol. 58 No. 4

21st Century Directions in Biology: Genomics Classifies Herps

Spring Spotlight on Books

$5.95 • $6.95 Canada

04

7 25274 05098 7

www.biosciencemag.org

Systematics and Herpetology in the Age of Genomics

JOHN J. WIENS

How is organismal biology changing in the era of genomics? Here, I discuss one example, the changes and trends in the systematics of reptiles and amphibians. The polymerase chain reaction, automated sequencing, and genomic tools now make it possible to apply a vast number of molecular characters to questions of phylogeny and species limits. At higher taxonomic levels, recent studies using these data have revealed some unexpected relationships, but also strong support for many traditionally recognized groups. At lower levels, molecular studies suggest that numerous species have been hidden by misleading taxonomy and morphological conservatism. However, the computational tools for analyzing multilocus data for phylogenetics and species delimitation are in need of further development, including greater integration with population genetics. Given current trends, much of reptile and amphibian phylogeny may soon be resolved. Although opportunities for tree-making by future systematists may shrink, opportunities for using phylogenies to address evolutionary and ecological questions should blossom.

Keywords: amphibians, genomics, phylogeny, reptiles, systematics

How are genomics and its associated molecular tools changing biology in the 21st century? Of course, this is impossible to answer in the space of a single article, because genomics has affected so many biological subdisciplines in so many ways. Instead, I present here a microcosm of these changes, focusing on a small sliver of the biological layer cake, both conceptually and organismally. I will discuss the impact of genomics on the systematics (phylogeny and species limits) of two groups of organisms that are familiar to everyone, reptiles (snakes, lizards, turtles, tuatara, crocodilians) and amphibians (frogs, toads, salamanders, caecilians). The study of these two groups is called herpetology. I will address how genomics is changing the practice of herpetological systematics, what we have learned from these new data so far, and what the future might hold.

Although systematics may seem like an abstruse biological subdiscipline to some, it is important to remember that most biological research depends on systematics at some level. For example, it is through systematics that species are discovered, described, and given scientific names. The study of phylogeny tells us fundamentally what organisms are (e.g., a dolphin is more closely related to a human than to a tuna), and allows us to make inferences about how their traits have evolved over time.

A brief history of molecular data in herpetological systematics

To understand how these new sources of molecular data are changing herpetological systematics, we need a historical context. Before the widespread use of DNA (deoxyribonucleic acid) sequencing, most molecular studies utilized data from allozymes and albumin immunological distances. Allozyme data typically consist of the frequencies of different alleles at a given enzyme locus, where these alleles are detected on the basis of their different mobilities in a starch or acetate gel. In the 1980s, allozyme data became widely used in phylogenetic studies of closely related reptile and amphibian species (e.g., Hillis et al. 1983). Allozyme studies also led to the discovery of new species that were previously unrecognized because of their morphological similarity (e.g., Highton et al. 1989). However, these data are of limited use for higher-level phylogeny, mostly because distantly related species tend to have no alleles in common. Allozymes are also problematic in that the underlying data (mobility of alleles) are entirely relative, and so raw data from different studies generally cannot be directly compared or combined.

Albumin immunological data were also used in many herpetological studies in the 1970s and 1980s (e.g., Maxson and Wilson 1974). Immunological data are based on overall similarity between molecules rather than homology of individual characters. Perhaps because of this, immunological data and the resulting phylogenies have not been widely embraced in herpetology. These data also suffer from the same problems of allozyme data, in terms of being entirely relative and difficult to apply over larger phylogenetic scales.

John J. Wiens (e-mail: wiensj@life.bio.sunysb.edu) is with the Department of Ecology and Evolution at Stony Brook University in Stony Brook, New York.

In the 1990s, DNA sequence data became widespread in systematic studies, thanks in large part to polymerase chain reaction (PCR) amplification of targeted sequences and the increasing automation and decreasing cost of DNA sequencing. The use of "universal" primers made it possible to amplify and sequence a targeted fragment of DNA in almost any reptile or amphibian species (e.g., Kocher et al. 1989). DNA sequence data are potentially relevant to any timescale (although some genes clearly are better for some timescales than for others), and data for the same gene from different studies are often easily combined. Combining data from different studies has also been greatly facilitated by GenBank, an online public database for sequence data (most journals now require that sequences be deposited there before publication).

Most early studies focused on mitochondrial genes (e.g., Hedges et al. 1991, Moritz et al. 1992, Arevalo et al. 1994) with only a few exceptions (i.e., nuclear ribosomal sequences; Larson and Dimmick 1993). Even today, it seems likely that the majority of DNA sequence studies published for reptiles and amphibians thus far are based on mitochondrial DNA data.

But mitochondrial DNA data have both advantages and disadvantages for systematics. On the plus side, mitochondrial genes are relatively easy to amplify and sequence. The mitochondrial genome also has a relatively fast mutation rate in vertebrates, providing an abundance of potentially informative variation, even among closely related species and conspecific populations (Avise 2000). Furthermore, the mitochondrial genome has a smaller effective population size than the typical nuclear gene (Avise 2000). Thus, the mitochondrial genome may be less subject to the problem of retained ancestral polymorphism (figure 1), and so, all other things being equal, the mitochondrial phylogeny may tend to track the phylogeny of the species better than a typical nuclear gene (Moore 1995).

On the negative side, the fast mutation rate of the mitochondrial genome can become a serious disadvantage at deeper phylogenetic levels. When rates of mutation are high, relatively long branches in a phylogeny (those expected to have accumulated many changes) may tend to be erroneously placed together by many phylogenetic methods because they will accumulate many shared traits by chance alone (Felsenstein 1978, 2004, Huelsenbeck 1995). This phenomenon is known as long-branch attraction. The problem is particularly disturbing because the incorrect results may have strong statistical support (e.g., from bootstrapping), and adding more fast-evolving characters may only exacerbate the problem (Felsenstein 1978, 2004). This problem may even extend down to the level of relatively closely related genera within a family, at least for some fast-evolving mitochondrial genes (e.g., cytochrome *b* and ND4 in iguanas; Wiens and Hollingsworth 2000).

In addition, the mitochondrial genome is inherited as a single unit (Avise 2000). This means that processes that may mislead phylogenies based on a single gene may extend to every gene in the mitochondrial genome. For example, introgression (hybridization) can occur between species that are not each other's closest relatives, causing a gene from one species to appear in the genome of another and causing a phylogeny based on this gene to incorrectly show these distantly related species to be sister species. When mitochondria introgress, every single gene in the mitochondrial genome will tend to suggest this same misleading pattern. Potential cases of this phenomenon are beginning to accumulate in the herpetological literature (e.g., Leaché and McGuire 2006). In summary, although mitochondrial data may work very well in many cases, to be certain that the relationships have been correctly inferred, it is invaluable to have other types of data as well.

Nuclear genomic data: Opening the floodgates

In the past few years, data from nuclear loci have begun to appear commonly in phylogenetic studies of reptiles and amphibians. Most studies so far have focused on a limited number of genes, including RAG-1, *c-mos,* and *c-myc.* But now, the number of genes that can be applied to a given study is essentially unlimited (or at least there are more genes available than most investigators will have the time and money to sequence).

This increase in the number of potentially usable genes has several causes. One critical factor is the pervasiveness of homology. For example, 75% of the genes in the human genome seemingly are homologous with those in the pufferfish (*Fugu rubripes*), with a total of approximately 28,000 shared genes (Aparicio et al. 2002). These two species represent clades that bracket amphibians and reptiles, such that most of these genes should also be shared by all reptiles and amphibians. Furthermore, nuclear genes typically are relatively slow evolving, so that primers that work well in one vertebrate clade often work well in others (e.g., primer pairs designed for amplifying genes in lizards can also give excellent results in frogs; Smith et al. 2007, Townsend et al. 2008). GenBank, and the associated computational tools that facilitate searching this database, is also tremendously important, making the sequences of thousands of genes for thousands of species freely available to anyone with an Internet connection. In many ways, the hardest part in finding new genes now is not that there are too few genes to choose from, but rather that the sheer number of genes available is almost overwhelming. Fortunately, studies are beginning to mine these genomic resources to identify a more restricted (but still very large) number of nuclear genes that are useful for vertebrate phylogenetics (e.g., Li et al. 2007, Townsend et al. 2008).

The increasing number of potential genes has been paralleled by other important trends in systematics. First, the taxonomic scale of studies is also (generally) increasing, thanks in part to the increasing ease and decreasing cost of automated DNA sequencing. Thus, researchers are collecting data not only from thousands of characters but also from dozens or even hundreds of taxa (e.g., Bossuyt et al. 2006, Frost et al. 2006). Although there has been extensive debate over the relative importance of sampling taxa or characters (e.g., Graybeal 1998,

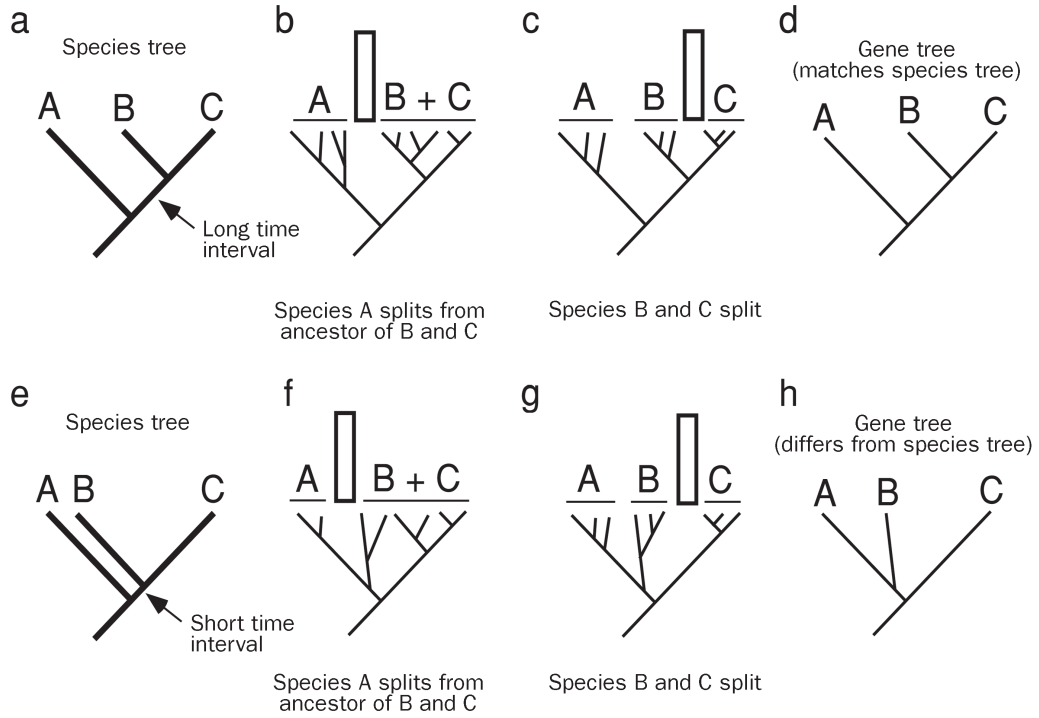

Figure 1. The phylogeny of a gene can differ from the phylogeny of species. This is especially likely when the time intervals between speciation events are short (and effective population sizes are large), leading to incomplete sorting of ancestral polymorphisms among lineages. This pair of hypothetical examples is intended to illustrate this idea, contrasting scenarios without (a–d) and with (e–h) incomplete lineage sorting. In (a)–(d), the time interval between the splitting events (i.e., between species A and the ancestor of species B + C and between species B and C) is relatively long. In (e)–(h) , this time interval is short. In this latter case, when the split between A and the B + C ancestor occurs, some of the alleles in the new species (B + C) are still more closely related to species A than they are to other individuals in species B + C. Given enough time, these anomalous alleles would disappear through genetic drift. However, when the B + C split occurs very soon afterward, these alleles are retained, and become fixed in species B by chance. In this case, the phylogeny of the gene differs from the phylogeny of the species. Note that when branches are very short, many different alternate gene trees could be generated through this process, and so gene trees may tend to disagree extensively. In b, c, f, and g, the terminal branches of the trees represent individual alleles from different populations, the vertical bar represents a geographic barrier between populations, and the horizontal line represents sets of interbreeding populations (species).

Rosenberg and Kumar 2001), researchers are no longer forced to choose between sampling large numbers of characters and large numbers of taxa for a given study. They can do both.

Perhaps just as important as the advances in data acquisition have been advances in data analysis. For example, methods have been developed that allow one to estimate the models underlying the evolution of DNA sequences, and then apply that information to phylogeny reconstruction (e.g., likelihood, Bayesian methods; Felsenstein 2004). The accuracy of these methods has been tested extensively with computer simulations (e.g., Huelsenbeck 1995, Wilcox et al. 2002, Alfaro et al. 2003). New algorithms and increasing computational power now make it possible to effectively analyze enormous data sets in a relatively short amount of time, even when using sophisticated model-based methods, thousands of species, and tens of thousands of characters (e.g., Stamatakis 2006).

The availability of vast numbers of nuclear loci does not mean that every phylogenetic problem will be solved easily, however. For example, finding a large number of nuclear genes that are evolving at the appropriate rate for a given phylogenetic problem can still be challenging, especially for studies at lower taxonomic levels that require rapidly evolving genes. Part of the problem is that rapidly evolving genes are less likely to have conserved primer sites, and so can be difficult to amplify and sequence.

Nuclear introns offer one potential solution. Introns are noncoding and thus free to evolve rapidly, but are flanked by

exons, which are more conserved (thus, primers can be designed that target exons flanking a desired intron). Many primers for nuclear introns are now available that are potentially usable across vertebrates (e.g., Lyons et al. 1997, Friesen et al. 1999, Dolman and Phillips 2004). However some significant problems still remain. First, whether a given locus will work in a particular clade still seems to be quite hit-or-miss (e.g., because of variation in intron length among clades, and other factors). Second, even fast-evolving introns may offer limited information in very recent or slow-evolving groups (e.g., turtles). Third, nuclear genes may often retain ancestral polymorphisms that are shared among closely related species (figure 1), which can confuse attempts to reconstruct phylogenies and species limits (although methods are now being developed that can potentially overcome this problem by explicitly considering population-genetic processes; Maddison and Knowles 2006, Knowles and Carstens 2007). The first two problems may also be ameliorated somewhat by using methods geared toward obtaining large numbers of nuclear markers that vary among closely related species (e.g., anonymous nuclear markers from whole genomic DNA, Carstens and Knowles 2006; SNPs [single nucleotide polymorphisms] from whole or partial genomes, Shaffer and Thomson 2007).

A more fundamental problem comes in the analysis of data from multiple nuclear loci. It is now becoming relatively straightforward to obtain data from, say, 10 nuclear loci for a given set of 20 species. But how exactly does one take these data and make a phylogeny? Whether or not data from different genes should be combined for phylogenetic analysis was a major debate in systematics in the 1990s (review in de Queiroz et al. [1995]). Although combined analysis has become standard, this issue has recently reemerged with a vengeance. Some recent studies (e.g., Poe and Chubb 2004, Rokas and Carroll 2006, Wiens et al. 2008) suggest that for very short internodes on phylogenies, there may be little agreement among the underlying gene trees because of the problem of incomplete lineage sorting mentioned above (figure 1). In these cases, combined analyses of multilocus data may be unsuccessful or might even be positively misled (e.g., Degnan and Rosenberg 2006, Kubatko and Degnan 2007). Methods are now being developed that may overcome this problem, again by incorporating information on population-genetic (coalescent) processes (e.g., Edwards et al. 2007). Overall, optimal methods for analyzing multilocus nuclear data are lagging behind the acquisition of these data, but the solution may lie in a greater integration of the fields of systematics and population genetics.

What have DNA data taught us about the evolution of reptiles and amphibians?

New DNA sequence data, especially from slow-evolving nuclear loci developed using genomic resources, are revolutionizing and refining our understanding of the evolution of many groups of reptiles and amphibians. Remarkably, most of these discoveries have appeared only within the past four years. I will discuss some of the major findings group by group.

Salamanders. Data from nuclear genes have helped provide a strongly supported hypothesis for salamander phylogeny at the family level (figure 2; Wiens et al. 2005, Roelants et al. 2007, Wiens 2007). This phylogeny shows some important differences from hypotheses based primarily on morphology (Gao and Shubin 2001) and mitochondrial DNA sequences (Weisrock et al. 2005). Paedomorphosis, the retention of larval or juvenile features of the ancestors in the adult stage of the descendants, appears to be a major problem for salamander phylogenetics based on morphology. Analyses of morphological data alone reconstruct clades of paedomorphic species that are seemingly both wrong and statistically well supported (Wiens et al. 2005). Conversely, long-branch attraction appears to be highly problematic in analyses of salamander phylogeny based on mitochondrial DNA (Weisrock et al. 2005).

New molecular data have also led to a major upheaval within the family Plethodontidae, which contains the majority of salamander species. Although traditional taxonomy recognized the subfamilies Desmognathinae and Plethodontinae, new molecular data (e.g., Chippindale et al. 2004, Mueller et al. 2004) suggest that Desmognathinae is actually nested deep inside of Plethodontinae. The implications of this finding go beyond merely shuffling names. The new phylogeny (along with detailed statistical analyses) strongly suggests that those desmognathine species with a larval stage actually evolved from ancestors with direct development, and that the larval stage has reevolved (Chippindale et al. 2004). Furthermore, new phylogenies within plethodontids suggest that the remarkable projectile tongue system of bolitoglossine plethodontids actually evolved twice, given the recent finding that *Hydromantes* is only distantly related to other bolitoglossines (Mueller et al. 2004).

Frogs. Major features of frog phylogeny (figure 2) have now been clarified by several studies incorporating slow-evolving nuclear genes (e.g., Hoegg et al. 2004, Roelants and Bossuyt 2005, Bossuyt et al. 2006, Roelants et al. 2007, Wiens 2007). Many traditionally recognized families and clades are now confirmed as monophyletic by DNA data, including Pipoidea, Neobatrachia, Hyloidea (Bufonoidea), and Ranoidea, but with some surprising exceptions. The relationships of the hyloid frogs are still in need of considerable attention, however, especially given the gross polyphyly of the most species-rich family of frogs (Leptodactylidae; Frost et al. 2006). Unfortunately, they may also prove to be very difficult to resolve, given the very short branches separating many of the major hyloid lineages (Roelants et al. 2007, Wiens 2007).

Caecilians. Caecilians are a poorly known and species-poor (approximately 175) group of tropical, burrowing, limbless amphibians. Recent studies of higher-level caecilian phylogeny based on multiple nuclear loci have now yielded a strongly supported hypothesis of caecilian relationships

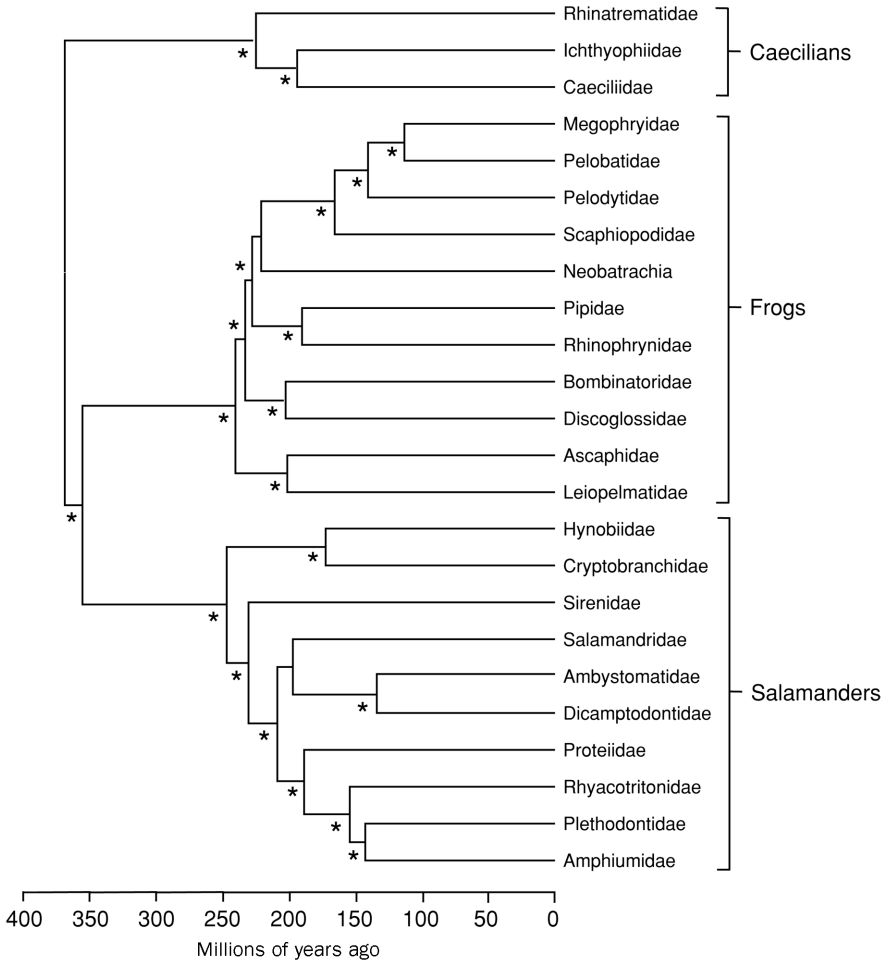

Figure 2. Phylogeny of the major groups of amphibians, based on maximum likelihood and Bayesian analysis of one mitochondrial gene and four nuclear genes (see figure 1 in Roelants et al. [2007]). Rather than using arbitrary branch lengths, the branch lengths depict divergence dates of clades estimated by a "relaxed" molecular clock method that incorporates multiple fossil calibration points (Roelants et al. 2007). Asterisks indicate strong support for clades (Bayesian posterior probabilities ≥ 0.95).

(figure 2) that is generally concordant with hypotheses based on morphology (Roelants et al. 2007).

"The" amphibian Tree of Life? A recent study has claimed to offer "the" amphibian Tree of Life (Frost et al. 2006), primarily on the basis of molecular data, along with an extensively revised classification of amphibians to the species level. Although many herpetologists are starting to use this classification, there is a need for some caution regarding both the tree and the taxonomy. For example, although this analysis included approximately 2400 characters of mitochondrial DNA data and approximately 2300 characters of nuclear data, the data matrix contains 15,320 characters. This unexplained discrepancy implies that the majority of the characters were artifacts of alignment. The tree includes several relationships that are rejected by model-based analyses of nuclear data (e.g., Roelants et al. 2007, Wiens 2007), but which appear in analyses of mitochondrial data alone and seem to be artifacts of long-branch attraction (e.g., sister relationship between proteids and sirenids; Weisrock et al. 2005). But perhaps the strongest evidence that this taxonomy may have been somewhat premature comes from the fact that of the 42 anuran families recognized by Frost and colleagues in 2006, Frost's Web site of amphibian classification in 2007 (*http://research.amnh.org/herpetology/amphibia/index.php*) no longer recognizes two of these families and has subdivided others to create five additional families. At this rate, much of the taxonomy Frost and colleagues proposed (2006) will be overturned by these same authors within just a few years.

Squamates. Molecular data have also led to a remarkable and unanticipated new hypothesis of squamate (lizard and snake) phylogeny (figure 3). Traditional phylogenies based on morphological characters (e.g., Estes et al. 1988) divided squamates into iguanians (iguanids, agamids, chameleons) and scleroglossans (geckoes, skinks, and all other lizards, as well as snakes and amphisbaenians). A new phylogeny incorporating slow-evolving nuclear loci (Townsend et al. 2004, Fry et al. 2006) suggests that iguanians form a clade with snakes and anguimorphans (i.e., alligator, glass, and monitor lizards, Gila monsters, and their relatives), and that this clade is nested deep within Scleroglossa. This new phylogeny also resolves the phylogenetic placement of several limb-reduced, burrowing lineages that were uncertain on the basis of morphological evidence (Estes et al. 1988), including the amphisbaenians, dibamids, and snakes (Townsend et al. 2004).

Recent studies have also addressed higher-level snake phylogeny (e.g., Slowinski and Lawson 2002, Wilcox et al. 2002, Lawson et al. 2005, Vidal et al. 2007a, Wiens et al. 2008), and are supporting many traditionally hypothesized aspects of the tree (i.e., basal scolecophidians, derived colubrids, viperids, and elapids; and the boas, pythons, and various burrowing lineages in the middle). However, members of some families (and other groups) have been found to be only distantly related to each other (e.g., Tropidophiidae). Another surprising result of recent studies is that Colubroidea, a relatively recent clade that contains the majority of snake species and seemingly represents a rapid radiation, so far appears to be

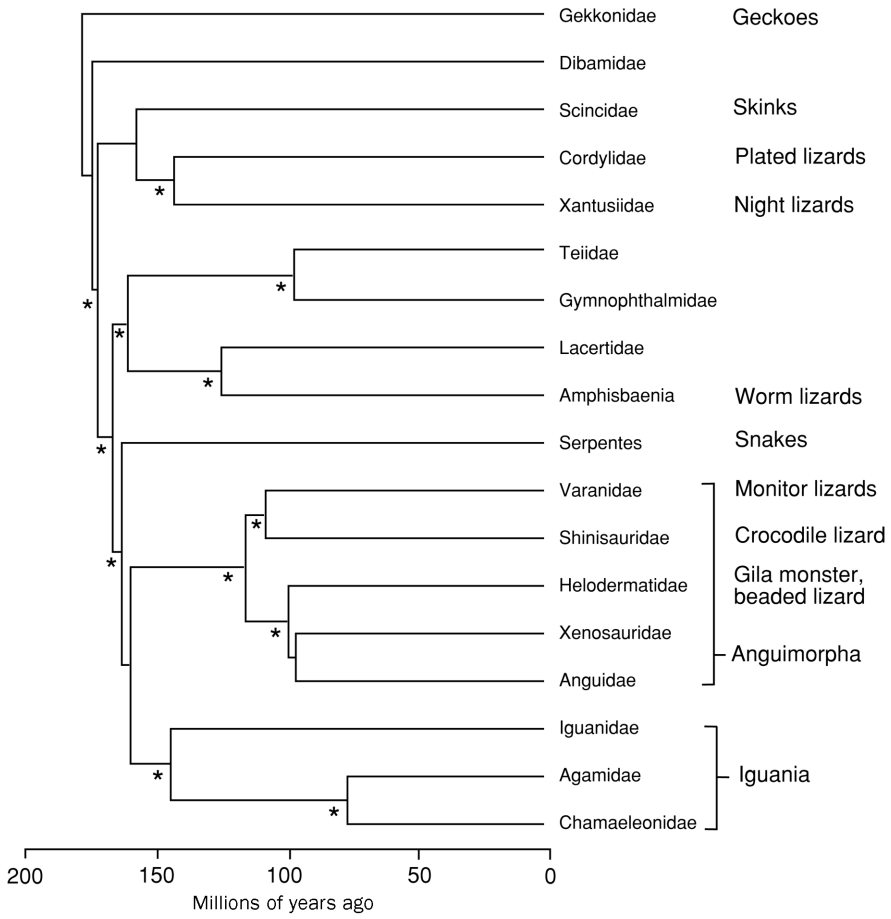

Figure 3. Phylogeny of squamate reptiles (lizards and snakes) based on maximum likelihood and Bayesian analysis of nuclear and mitochondrial genes (see figure 7b in Townsend et al. [2004]). Branch lengths reflect estimated divergence times from Wiens and colleagues (2006a), using a "relaxed" molecular clock method that incorporates multiple fossil calibration points (see also Hugall et al. [2007]). Asterisks indicate strong support for clades (Bayesian posterior probabilities ≥ 0.95).

relatively easy to resolve on the basis of molecular data (e.g., Lawson et al. 2005, Vidal et al. 2007b, Wiens et al 2008).

Turtles. Recent analyses of turtle phylogeny based on the nuclear RAG-1 gene and mitochondrial sequences (figure 4; Krenz et al. 2005) support a tree that shares many clades with those based on morphology and traditional taxonomy (e.g., Gaffney and Meylan 1988), with some exceptions. In many ways, turtle phylogeny now appears to be relatively well resolved.

However, the phylogenetic position of turtles within reptiles has been more surprising. Although traditional hypotheses place turtles as the sister group of lepidosaurs (tuatara, lizards, snakes) + archosaurs (crocodilians + birds), recent analyses based on nuclear and mitochondrial loci (e.g., Hedges and Poling 1999, Iwabe et al. 2005, Hugall et al. 2007) suggest that turtles are instead the sister group to extant archosaurs.

Crocodilians. The phylogeny of crocodilians (figure 5) offers a bizarre case of concordance and discordance between molecular and morphological data (e.g., Poe 1996, Harshman et al. 2003, Gatesy et al. 2003). These data give largely concordant results, with the exception of one species, the gavial (*Gavialis*). The morphological data strongly place *Gavialis* as the sister group to all other extant crocodilians, despite its superficial morphological similarity to the false gavial (*Tomistoma*). Yet, the molecular data (and combined analyses of the molecular and morphological data) place the true gavial with the false gavial (*Tomistoma*) as sister group to other Crocodylidae.

Morphology versus molecules: No easy answers

Many molecular studies are suggesting relationships that differ radically from previous morphological studies. This raises the obvious question: which hypothesis is correct? This question must be addressed on a case-by-case basis. In many cases, the morphological hypotheses were only weakly supported, and the incongruence may simply reflect the morphological data being inconclusive. For some groups, the most widely used morphology-based phylogenies were not based on an actual analysis (e.g., squamates, Estes et al. 1988; anurans, Ford and Cannatella 1993). In other cases, there may be a relatively clear source of error in the molecular data (e.g., long-branch attraction, introgression) or in the morphological data (e.g., paedomorphosis). However, in some cases of incongruence, the results are statistically well supported by both types of data and there is no obvious explanation. These include the conflicts over the position of *Gavialis* in crocodilians and of Iguania within squamates.

Although I have emphasized the conflicts between molecular and morphological results, congruence between them appears to be far more widespread. For example, even though molecular and morphological hypotheses for squamate phylogeny are strikingly discordant, an analysis of squamate phylogeny on the basis of nuclear RAG-1 data nevertheless supports 88% of the morphology-based families that were represented by two or more species (see figure 1 in Townsend et al. [2004]).

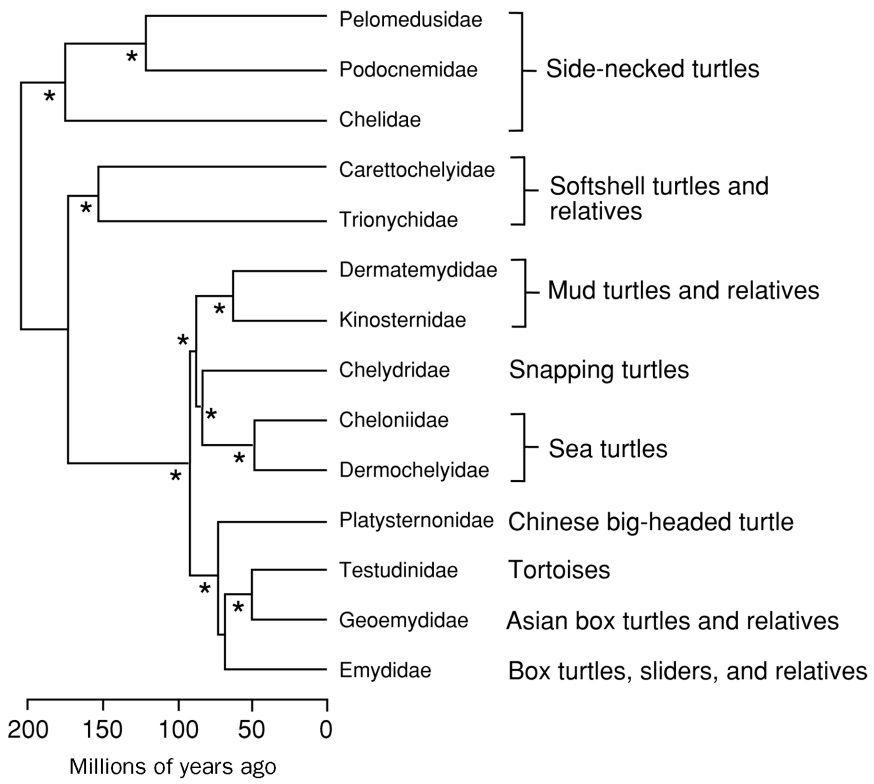

Figure 4. Phylogeny of turtles based on maximum likelihood and Bayesian analysis of nuclear and mitochondrial genes (see figure 5b in Krenz et al. [2005]). Branch lengths reflect estimated divergence times from Near and colleagues (2005), using a "relaxed" molecular clock method that incorporates multiple fossil calibration points (but note that there is a slight difference in the trees from Krenz and colleagues [2005] and Near and colleagues [2005] regarding placement of* Chelydra*). Asterisks indicate strong support for clades (Bayesian posterior probabilities ≥ 0.95).

Despite the increasing prevalence of molecular data and molecular systematists, phylogenetic analyses of morphology remain critically important. Although molecular data offer a much larger number of characters than morphological data sets, molecular data (so far) cannot directly address the phylogenetic placement of fossil taxa, and we still have a poor understanding of the phylogenetic relationships of many fossil reptiles and amphibians. Correctly placing these fossil taxa is critical for a complete understanding of the phylogeny, biogeography, and evolution of reptiles and amphibians, and for calibrating molecular clock analyses to estimate the divergence times of living taxa (e.g., figures 2–4).

Hidden diversity at the species level

Molecular systematic studies have also revealed surprising results at the species level. Again and again, studies have found that what appeared to be a single species actually consists of two or more lineages that are as divergent as traditionally recognized species. Thus, current taxonomy appears to be hiding a great deal of species diversity, which is now being revealed by DNA studies.

This diversity has been hidden in two distinct ways, depending somewhat upon the taxonomic group. In reptiles, distinct species are sometimes hidden more by taxonomy than by morphology. Many reptile species were previously divided into subspecies, usually on the basis of geographic variation in scalation and coloration. Molecular studies are revealing that many subspecies seem to represent distinct species (e.g., Zamudio et al. 1997, Ashton and de Queiroz 2001). Thus, the diversity was recognized morphologically, but was classified inappropriately. However, there is not always a perfect correspondence between subspecies and species limits, and many studies find that a single polytypic species contains some subspecies that are distinct at the species level and others that are completely meaningless.

In contrast, species diversity in amphibians seems to be hidden more often by morphological conservatism than by incorrect ranking of subspecies. Most amphibian species were never divided into subspecies, most likely because they have fewer morphological characters that vary geographically (e.g., most lack scales). Instead, many amphibian species that lack obvious geographic variation in morphology appear to consist of multiple species that are distinct at the molecular level (e.g., Kozak et al. 2005, Stuart et al. 2006). This morphological conservatism has been commented on since the days of allozymes (e.g., Larson 1984), and is being confirmed again and again by DNA-based studies.

Unfortunately, much of the cryptic diversity that has been recently discovered through analyses of DNA data has been based exclusively on mitochondrial genes. Perhaps because of this, dozens of potential species have been discovered but not officially described, and remain in a taxonomic and conservation limbo. Their status should be resolved using additional evidence (e.g., nuclear introns) so that they can be formally named. Determining species limits from DNA data is also an area where data acquisition has outstripped the methods for data analysis, and where new methods are being developed to meet this challenge (e.g., Pons et al. 2006, Knowles and Carstens 2007). This is also an area that would benefit from greater integration of systematics and population genetics (e.g., Knowles and Carstens 2007). After all, species delimitation hinges on inferred patterns of gene flow,

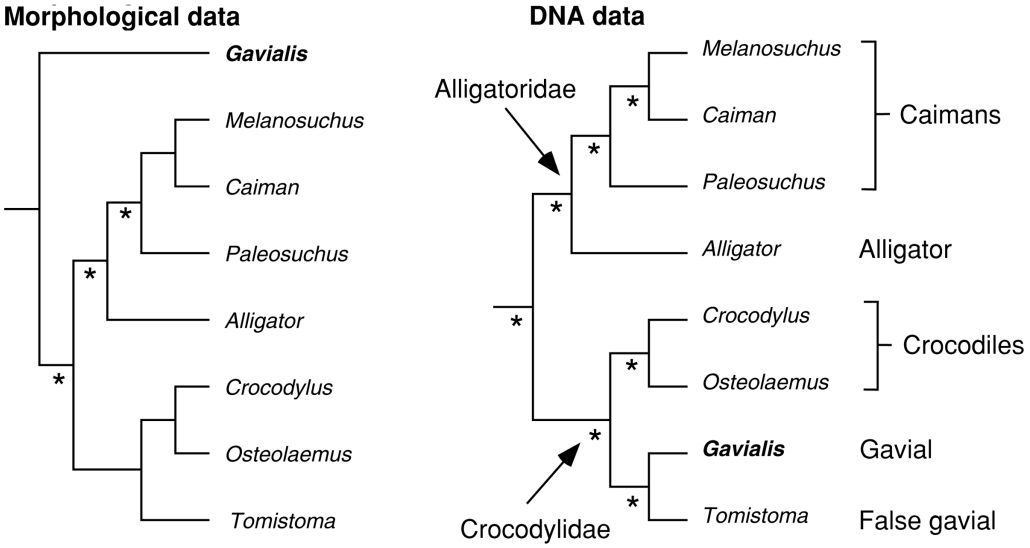

Figure 5. Phylogeny of crocodilian reptiles based on parsimony analyses of morphology and nuclear and mitochondrial genes (see figure 2 in Gatesy et al. [2003]). Branch lengths are arbitrary. Asterisks indicate strong support for clades (parsimony bootstrap values ≥ 95%).

is complicated by genetic drift (e.g., figure 1), and often uses patterns of gene coalescence.

The future of herpetological systematics

The study of reptile and amphibian systematics seems to be in the middle of a golden age. Tools such as PCR, automated sequencing, and genomics now provide an almost limitless source of characters that can be used to resolve phylogeny and species limits. These tools are now becoming extensively used, and major projects are under way on the phylogeny of amphibians, squamates, and many species-rich groups within Squamata (e.g., geckoes). Relatively speaking, crocodilians, turtles, and caecilians have so few species that obtaining a complete phylogeny of all extant species for these groups seems almost trivial. It should also be noted that the ambitious scale of many of these studies is also being facilitated by extensive collaboration among researchers (both nationally and internationally), seemingly part of a larger trend within the sciences.

But when I say that we are in the *middle* of a golden age, I also mean that the end may be in sight. Given the scale and success of recent studies, it seems likely that the phylogeny of all major groups will be resolved at the level of families and genera within the next 10 years (if not sooner; figures 2–5).

What does this leave for the next generation of systematic herpetologists? Although most higher level phylogeny may soon be resolved, filling in the final twigs of the reptile and amphibian trees will go on for much, much longer. This must be true because new species of reptiles and amphibians continue to be described at a rapid rate (e.g., a 25% increase in amphibian species over the previous 11 years was recently documented; Köhler et al. 2005). Furthermore, the proportion of wide-ranging species that have been the subject of detailed phylogeographic studies is still relatively small, and most were based on only a single genetic locus (i.e., mitochondrial DNA).

In addition, even though all families and most genera will be included in phylogenetic trees in the near future, it seems safe to say that not every branch will be strongly supported (e.g., figures 2–4). Some clades may prove very difficult to resolve. For example, a study of higher-level snake phylogeny found that combined analysis of 20 nuclear genes (13,322 characters) offers only weak support for 31% of the 48 clades (Wiens et al. 2008). Most of these clades that remain poorly resolved are associated with very short branches, where there are often strongly supported conflicts among genes and similar numbers of genes that strongly support alternate topologies. If and how such very short branches can be resolved is still an open question (e.g., Poe and Chubb 2004, Rokas and Carroll 2006, Edwards et al. 2007). Nevertheless, opportunities to discover something truly novel about the phylogeny of major groups will most likely become increasingly rare for students of herpetology in the future.

Does this mean that we should stop training herpetological systematists? For several reasons I would say both "no" and "yes." I would argue that we should continue training students with expertise in systematics, but not train them to be merely systematists. Every group of organisms needs biologists who know the diversity, distribution, and natural history of its species. Furthermore, given the accelerating biodiversity crisis (which is particularly grave for amphibians; Stuart et al. 2004), we must keep describing new species and placing these new species into the Tree of Life. But within reptiles and amphibians, the major phylogenetic questions may soon either be solved or found to be insoluble (this may make reptiles and amphibians very different from most other groups of organisms, but perhaps not so unusual for vertebrates).

One way that students might deal with this issue is to focus on using phylogenies in addition to making them. Phylogenies are necessary to address evolutionary and ecological questions that involve the evolution of traits among species and populations (e.g., Felsenstein 1985, Donoghue 1989). Using phylogenies requires knowledge and skills that presently are associated mainly with systematists, including the diversity, natural history, distribution, and identification of species within a group, as well as methods for divergence date estimation, ancestral trait reconstruction, and phylogeny-based

analysis of correlation between traits. Analyses that use phylogenies to address broader questions are also being invigorated by the new flood of molecular data and the associated computational tools. For example, to address the evolutionary and ecological causes of the latitudinal diversity gradient, we were able to utilize a molecular phylogeny for 325 frog species (Wiens et al. 2006b). In summary, I think herpetologists should continue to train systematists, but we might do well to diversify their training to go beyond systematics alone.

Conclusions

New molecular data are transforming the field of herpetology in a number of ways, and many of these same transformations have likely occurred or may be occurring in other groups of organisms. PCR, automated sequencing, and genomics have created a virtually inexhaustible source of new characters for systematics. This new wealth of data has been paralleled by important advances in computational phylogenetic methods. Together, these molecular and computational tools are overturning or questioning many traditional ideas about reptile and amphibian phylogeny based on morphology. In some cases, this is because these previous hypotheses were only weakly supported by the morphological data. But there are other cases of incongruence that have interesting biological causes (e.g., paedomorphosis), or where the causes of incongruence remain somewhat mysterious. Molecular studies are also revealing hidden diversity within many reptile and amphibian species, suggesting that the true species richness of reptiles and amphibians will not be understood without detailed molecular studies across the geographic ranges of many species.

Yet, despite this new flood of data and many computational advances, the most appropriate analytical methods for using multilocus data in systematics are uncertain and are still being actively developed and tested. This is true both for determining species limits and resolving the phylogeny among species. Methods for addressing both questions will increasingly utilize information from the field of population genetics (e.g., coalescence theory; Maddison and Knowles 2006, Edwards et al. 2007, Knowles and Carstens 2007).

Given current trends, it seems likely that the phylogeny of most reptile and amphibian groups will be resolved in the next 10 years, at least at the level of currently recognized genera. Of course, a comprehensive tree that includes all species and has strong support for every branch will take much, much longer, especially given the rapid rate at which new reptile and amphibian species continue to be discovered and the difficulty of resolving very short branches. Nevertheless, new students should consider broadening their research questions beyond organismal phylogeny and species limits, but while still maintaining the knowledge and skills of systematics and herpetology.

Acknowledgments

I am grateful to three anonymous reviewers for helpful comments. I thank the National Science Foundation for financial support of my research (most recently, EF 0334923).

References cited

Alfaro ME, Zoller S, Lutzoni F. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Molecular Biology and Evolution 20: 255–266.

Aparicio S, et al. 2002. Whole genome shotgun assembly and analysis of the genome of *Fugu rubripes.* Science 297: 1301–1310.

Arevalo E, Davis SK, Sites JW. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Systematic Biology 43: 387–418.

Ashton KG, de Queiroz A. 2001. Molecular systematics of the western rattlesnake, *Crotalus viridis* (Viperidae), with comments on the utility of the D-loop in phylogenetic studies of snakes. Molecular Phylogenetics and Evolution 21: 176–189.

Avise JC. 2000. Phylogeography: The History and Formation of Species. Cambridge (MA): Harvard University Press.

Bossuyt F, Brown RM, Hillis DM, Cannatella DC, Milinkovitch MC. 2006. Phylogeny and biogeography of a cosmopolitan frog radiation: Late Cretaceous diversification resulted in continent-scale endemism in the family Ranidae. Systematic Biology 55: 579–594.

Carstens BC, Knowles LL. 2006. Variable nuclear markers for *Melanoplus oregonensis* identified from the screening of a genomic library. Molecular Ecology Notes 6: 683–685.

Chippindale PT, Bonett RM, Baldwin AS, Wiens JJ. 2004. Phylogenetic evidence for a major reversal in life history evolution in plethodontid salamanders. Evolution 58: 2809–2822.

Degnan JH, Rosenberg NA. 2006. Discordance of species trees with their most likely gene trees. Public Library of Science Genetics 2: 762–768.

de Queiroz A, Donoghue MJ, Kim J. 1995. Separate versus combined analysis of phylogenetic evidence. Annual Review of Ecology and Systematics 26: 657–681.

Dolman G, Phillips B. 2004. Single copy nuclear DNA markers characterized for comparative phylogeography in Australian wet tropics rainforest skinks. Molecular Ecology Notes 4: 185–187.

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

Edwards SV, Liu L, Pearl DK. 2007. High-resolution species trees without concatenation. Proceedings of the National Academy of Sciences 104: 5936–5941.

Estes R, de Queiroz K, Gauthier JA. 1988. Phylogenetic relationships within Squamata. Pages 119–281 in Estes R, Pregill G, eds. Phylogenetic Relationships of the Lizard Families. Stanford (CA): Stanford University Press.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Systematic Zoology 27: 401–410.

———. 1985. Phylogenies and the comparative method. American Naturalist 125: 1–15.

———. 2004. Inferring phylogenies. Sunderland (United Kingdom): Sinauer Associates.

Ford LS, Cannatella DC. 1993. The major clades of frogs. Herpetological Monographs 7: 94–117.

Friesen VL, Congdon BC, Kidd MG, Birt TP. 1999. PCR primers for the amplification of five nuclear introns in vertebrates. Molecular Ecology 8: 2147–2149.

Frost D, et al. 2006. The amphibian tree of life. Bulletin of the American Museum of Natural History 297: 1–370.

Fry BG, et al. 2006. Early evolution of the venom system in lizards and snakes. Nature 439: 584–588.

Gaffney ES, Meylan PA. 1988. A phylogeny of turtles. Pages 157–219 in Benton MJ, ed. The Phylogeny and Classification of the Tetrapods, vol. 1. Oxford (United Kingdom): Clarendon Press.

Gao K, Shubin N. 2001. Late Jurassic salamanders from northern China. Nature 410: 574–577.

Gatesy J, Amato G, Norell M, DeSalle R, Hayashi C. 2003. Combined support for wholesale taxic atavism in gavialine crocodylians. Systematic Biology 52: 403–422.

Graybeal A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Systematic Biology 47: 9–17.

Harshman J, Huddleston CJ, Bullock JP, Parsons TJ, Braun MJ. 2003. True and false gharials: A nuclear gene phylogeny of Crocodylia. Systematic Biology 52: 386–402.

Hedges SB, Poling LL. 1999. A molecular phylogeny of reptiles. Science 283: 998–1001.

Hedges SB, Bezy RL, Maxson LR. 1991. Phylogenetic relationships and biogeography of xantusiid lizards inferred from mitochondrial DNA sequences. Molecular Biology and Evolution 8: 767–780.

Highton RJ, Maha GC, Maxson LR. 1989. Biochemical evolution in the slimy salamanders of the *Plethodon glutinosus* complex in the eastern United States, pt. 1: Geographic protein variation. Illinois Biological Monographs 57: 1–78.

Hillis DM, Frost JS, Wright DA. 1983. Phylogeny and biogeography of the *Rana pipiens* complex: A biochemical evaluation. Systematic Zoology 32: 132–143.

Hoegg S, Vences M, Brinkmann H, Meyer A. 2004. Phylogeny and comparative substitution rates of frogs inferred from sequences of three nuclear genes. Molecular Biology and Evolution 21: 1188–1200.

Huelsenbeck JP. 1995. The performance of phylogenetic methods in simulation. Systematic Biology 44: 17–48.

Hugall AF, Foster R, Lee MSY. 2007. Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. Systematic Biology 56: 543–563.

Iwabe N, Hara Y, Kumazawa Y, Shibamoto K, Saito Y, Miyata T, Katoh K. 2005. Sister group relationship of turtles to the bird-crocodilian clade revealed by nuclear DNA-coded proteins. Molecular Biology and Evolution 22: 810–813.

Knowles LL, Carstens BC. 2007. Delimiting species without monophyletic gene trees. Systematic Biology 56: 887–895.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences 86: 6196–6200.

Köhler J, Vieites DR, Bonett RM, Hita García F, Glaw F, Steinke D, Vences M. 2005. New amphibians and global conservation: A boost in species discoveries in a highly endangered vertebrate group. BioScience 55: 693–696.

Kozak KH, Larson A, Bonett RM, Harmon LJ. 2005. Phylogenetic analysis of ecomorphological divergence, community structure, and diversification rates in dusky salamanders (Plethodontidae: *Desmognathus*). Evolution 59: 2000–2016.

Krenz JG, Naylor GJP, Shaffer HB, Janzen FJ. 2005. Molecular phylogenetics and evolution of turtles. Molecular Phylogenetics and Evolution 37: 178–191.

Kubatko LS, Degnan JH. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56: 17–24.

Larson A. 1984. Neontological inferences of evolutionary pattern and process in the salamander family Plethodontidae. Evolutionary Biology 17: 119–217.

Larson A, Dimmick WW. 1993. Phylogenetic relationships of the salamander families: An analysis of congruence among morphological and molecular characters. Herpetological Monographs 7: 77–93.

Lawson R, Slowinski JB, Crother BI, Burbrink FT. 2005. Phylogeny of the Colubroidea (Serpentes): New evidence from mitochondrial and nuclear genes. Molecular Phylogenetics and Evolution 37: 581–601.

Leaché AD, McGuire JA. 2006. Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: Evidence for a misleading mitochondrial gene tree. Molecular Phylogenetics and Evolution 39: 628–644.

Li C, Ortí G, Zhang G, Lu G. 2007. A practical approach to phylogenomics: The phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evolutionary Biology 7: 44.

Lyons LA, Laughlin TR, Copeland NG, Jenkins NA, Womack JE, O'Brien SJ. 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nature Genetics 15: 47–56.

Maddison WP, Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. Systematic Biology 55: 21–30.

Maxson LR, Wilson AC. 1974. Convergent morphological evolution detected by studying proteins of tree frogs in the *Hyla eximia* species group. Science 185: 66–68.

Moore WS. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. Evolution 49: 718–726.

Moritz C, Schneider CJ, Wake DB. 1992. Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. Systematic Biology 41: 273–291.

Mueller RL, Macey JR, Jaekel M, Wake DB, Boore JL. 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. Proceedings of the National Academy of Sciences 101: 13820–13825.

Near TJ, Meylan PA, Shaffer HB. 2005. Assessing concordance of fossil calibration points in molecular clock studies: An example using turtles. American Naturalist 165: 137–146.

Poe S. 1996. Data set incongruence and the phylogeny of crocodilians. Systematic Biology 45: 393–414.

Poe S, Chubb AL. 2004. Birds in a bush: Five genes indicate explosive evolution of avian orders. Evolution 58: 404–415.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55: 595–609.

Roelants K, Bossuyt F. 2005. Archaeobatrachian paraphyly and Pangean diversification of crown group frogs. Systematic Biology 54: 111–126.

Roelants K, Gower DJ, Wilkinson M, Loader SP, Biju SD, Guillaume K, Bossuyt F. 2007. Patterns of diversification in the history of modern amphibians. Proceedings of the National Academy of Sciences 104: 887–892.

Rokas A, Carroll SB. 2006. Bushes in the Tree of Life. Public Library of Science Biology 4: e352.

Rosenberg MS, Kumar S. 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. Proceedings of the National Academy of Sciences 98: 10751–10756.

Shaffer HB, Thomson RC. 2007. Delimiting species in recent radiations. Systematic Biology 56: 896–906.

Slowinski JB, Lawson R. 2002. Snake phylogeny: Evidence from nuclear and mitochondrial genes. Molecular Phylogenetics and Evolution 24: 194–202.

Smith SA, Arif S, Nieto Montes de Oca A, Wiens JJ. 2007. A phylogenetic hotspot for evolutionary novelty in Middle American treefrogs. Evolution 61: 2075–2085.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.

Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, Waller RW. 2004. Status and trends of amphibian declines and extinctions worldwide. Science 306: 1783–1786.

Stuart BL, Inger RF, Voris HK. 2006. High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. Biology Letters 2: 470–474.

Townsend T, Larson A, Louis EJ, Macey JR. 2004. Molecular phylogenetics of Squamata: The position of snakes, amphisbaenians, and dibamids, and the root of the squamate tree. Systematic Biology 53: 735–757.

Townsend TM, Alegre ER, Kelley ST, Wiens JJ, Reeder TW. 2008. Rapid development of multiple nuclear loci for phylogenetic analysis using

genomic resources: An example from squamate reptiles. Molecular Phylogenetics and Evolution. Forthcoming.

Vidal N, Delmas A-S, David P, Cruaud C, Couloux A, Hedges SB. 2007a. The phylogeny and classification of caenophidian snakes inferred from seven nuclear protein-coding genes. Comptes Rendus Biologies 330: 182–187.

Vidal N, Delmas A-S, Hedges SB. 2007b. The higher-level relationships of alethinophidian snakes inferred from seven nuclear and mitochondrial genes. Pages 27–33 in Henderson RW, Powell R, eds. Biology of the Boas and Pythons. Eagle Mountain (UT): Eagle Mountain.

Weisrock DW, Harmon LJ, Larson A. 2005. Resolving the deep phylogenetic relationships among salamander families: Analyses of mitochondrial and nuclear genomic data. Systematic Biology 54: 758–777.

Wiens JJ. 2007. Global patterns of species richness and diversification in amphibians. American Naturalist 170: S86–S106.

Wiens JJ, Hollingsworth BD. 2000. War of the iguanas: Conflicting molecular and morphological phylogenies and long-branch attraction in iguanid lizards. Systematic Biology 49: 143–159.

Wiens, JJ, Bonett RM, Chippindale PT. 2005. Ontogeny discombobulates phylogeny: Paedomorphosis and higher-level salamander phylogeny. Systematic Biology 54: 91–110.

Wiens JJ, Brandley MC, Reeder TW. 2006a. Why does a trait evolve multiple times within a clade? Repeated evolution of snake-like body form in squamate reptiles. Evolution 60: 123–141.

Wiens JJ, Graham CH, Moen DS, Smith SA, Reeder TW. 2006b. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: Treefrog trees unearth the roots of high tropical diversity. American Naturalist 168: 579–596.

Wiens JJ, Kuczynski CA, Smith SA, Mulcahy D, Sites JW, Townsend TM, Reeder TW. 2008. Branch lengths, support, and congruence: Testing the phylogenomic approach with 20 nuclear loci in snakes. Systematic Biology. Forthcoming.

Wilcox TP, Zwickl DJ, Heath TA, Hillis DM. 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. Molecular Phylogenetics and Evolution 25: 361–371.

Zamudio, KR, Jones KB, Ward RH. 1997. Molecular systematics of short-horned lizards: Biogeography and taxonomy of a widespread species complex. Systematic Biology 46: 284–305.

doi:10.1641/B580405
Include this information when citing this material.