

## Multiple Data Sets, High Homoplasy, and the Phylogeny of Softshell Turtles (*Testudines: Trionychidae*)

TAG N. ENGSTROM,<sup>1,3</sup> H. BRADLEY SHAFFER,<sup>1</sup> AND WILLIAM P. MCCORD<sup>2</sup>

<sup>1</sup>Center for Population Biology and Section of Evolution and Ecology, University of California, One Shields Avenue, Davis, CA 95616, USA

<sup>2</sup>East Fishkill Animal Hospital Hopewell Junction, NY 12533, USA

<sup>3</sup>(Current address) Department of Biological Sciences, California State University at Chico, Chico, CA 95929-0515, USA;  
E-mail: tengstrom@csuchico.edu

**Abstract.**—We present a phylogenetic hypothesis and novel, rank-free classification for all extant species of softshell turtles (*Testudines: Trionychidae*). Our data set included DNA sequence data from two mitochondrial protein-coding genes and a ~1-kb nuclear intron for 23 of 26 recognized species, and 59 previously published morphological characters for a complimentary set of 24 species. The combined data set provided complete taxonomic coverage for this globally distributed clade of turtles, with incomplete data for a few taxa. Although our taxonomic sampling is complete, most of the modern taxa are representatives of old and very divergent lineages. Thus, due to biological realities, our sampling consists of one or a few representatives of several ancient lineages across a relatively deep phylogenetic tree. Our analyses of the combined data set converge on a set of well-supported relationships, which is in accord with many aspects of traditional softshell systematics including the monophyly of the *Cyclanorbinae* and *Trionychinae*. However, our results conflict with other aspects of current taxonomy and indicate that most of the currently recognized tribes are not monophyletic. We use this strong estimate of the phylogeny of softshell turtles for two purposes: (1) as the basis for a novel rank-free classification, and (2) to retrospectively examine strategies for analyzing highly homoplasious mtDNA data in deep phylogenetic problems where increased taxon sampling is not an option. Weeded and weighted parsimony, and model-based techniques, generally improved the phylogenetic performance of highly homoplasious mtDNA sequences, but no single strategy completely mitigated the problems associated with these highly homoplasious data. Many deep nodes in the softshell turtle phylogeny were confidently recovered only after the addition of largely nonhomoplasious data from the nuclear intron. [Homoplasy; mitochondrial DNA; multiple data sets; rank-free classification; nuclear intron; phylogeny; *Trionychidae*.]

The use of DNA sequence data has become nearly ubiquitous in systematics (Hillis et al., 1996). Mitochondrial DNA (mtDNA) sequence data has been and continues to be particularly popular because the conserved gene order, lack of introns, and lack of recombination in the mitochondrial genome render the acquisition and analysis of mtDNA sequence data relatively easy compared with the more complex nuclear genome. The rapid rate of nucleotide substitution in the mitochondrial genome (Brown et al., 1979) provides a rich source of variable characters. However, this rapid rate of substitution, combined with at most four character states, a consistently strong base compositional bias, and functional constraints (Graybeal, 1993; Meyer, 1994) all contribute to potentially high levels of homoplasy in mtDNA, particularly for more divergent phylogenetic lineages. High homoplasy levels may lead mtDNA studies to spurious conclusions (Naylor and Brown, 1998; Garcia-Machado et al., 1999; Wiens and Hollingsworth, 2000), bringing into question the general utility of mitochondrial data for deep phylogenetic questions (Naylor and Brown, 1998; Matthee et al., 2001). However, the actual effects that high levels of homoplasy have on phylogenetic reconstruction are not clear. Although homoplasy does have the potential to obscure phylogenetic information (Sanderson and Hufford, 1996), several studies have found a positive relationship between level of homoplasy in a data set and the level of resolution in the phylogeny (Sanderson and Donoghue, 1996; Källersjö et al., 1998, 1999), implying that homoplasy may not be all bad. Further countering this lingering doubt about the utility of mtDNA

is the observation that many relatively deep problems in vertebrate phylogenetics have apparently benefited from mtDNA analyses (Shaffer et al., 1997; Zardoya and Meyer, 1998; Hedges and Poling, 1999; Mindell et al., 1999).

In contrast to mtDNA, nuclear protein-coding genes and introns tend to evolve more slowly (Prychitko and Moore, 1997, 2000; Groth and Barrowclough, 1999; Birks and Edwards, 2002), making them less prone to excessive homoplasy. Nuclear introns have the further advantage of being free from many of the evolutionary constraints imposed on protein-coding sequences, resulting in phylogenetic markers, which, in vertebrates, usually show little base compositional bias, relatively low transition-transversion ratio, and little among-site rate heterogeneity (Armstrong et al., 2001; Prychitko and Moore, 2003; Fujita et al., 2004). One disadvantage of nuclear DNA is that the same slow rate of evolution, which makes nuclear DNA less prone to homoplasy on long time scales, can also result in a lack of variation on shorter time scales (Birks and Edwards, 2002). The identification of near-universal primers for a few genes, including RAG1 (Greenhalgh et al., 1993; Groth and Barrowclough, 1999), *c-mos* (Saint et al., 1998), beta fibrinogen intron 7 (Prychitko and Moore, 1997, 2000, 2003), and other introns (Friesen et al., 1999; Fujita et al., 2004), has helped make nDNA sequences more accessible to the vertebrate phylogenetics community. However, the fact remains that for most taxa, nuclear data are still not as easily obtainable as mtDNA data and it is possible that nuclear DNA data may never be collected for the many

studies that have been done exclusively using mtDNA. Thus many molecular phylogenetic analyses are limited to mitochondrial data even when these data are likely to be compromised by high levels of homoplasy. This fact, coupled with the incredible diversity of mtDNA sequence data now available for virtually any phylogenetic question, make it desirable to develop strategies for analyzing data sets containing highly homoplasious data typical of mtDNA, both alone and in combination with nuclear DNA in a way that that incorporates the strengths and overcomes the weaknesses of each.

Here we present phylogenetic analyses of mitochondrial and nuclear DNA sequence data for softshell turtles (*Testudines: Trionychidae*) separately, combined with each other and with previously published morphological characters. Our molecular data consist of mitochondrial sequence data from two protein-coding genes and nuclear sequence data from a ~1-kb intron from the R35 neural transmitter gene (Friedel et al., 2001; Fujita et al., 2004) from 23 of 26 recognized species of softshell turtles. We chose to work with these two mitochondrial genes because previous studies have shown them to be useful in resolving deep relationships in turtle phylogenetics (Shaffer et al., 1997; Starkey, 1997) and because their frequent use in vertebrate systematics implies that observations on analytical strategies that we have applied here might be more broadly applicable. The morphological data set consists of 59 osteological characters collected by Meylan (1987) for a complimentary set of 24 trionychid species.

A primary objective of this study was to produce a well-supported phylogeny for all extant softshell turtles. Although neither the molecular nor the morphological data set alone contains data for every recognized species, the combined data set includes complete taxonomic coverage for this globally distributed clade of turtles, with incomplete data for a few taxa. There are potential disadvantages to including taxa with incomplete data; however, these problems are usually outweighed by the advantages of their inclusion (Wiens and Reeder, 1995; Wiens, 1998, 2003). In this case we feel that the advantage of obtaining a complete phylogenetic hypothesis at the species level for all softshell turtles outweighs potential disadvantages.

For many taxa, mtDNA data from one or two genes are all that is currently available for phylogenetic analyses, and for better or for worse these existing mtDNA data may be all that is ever available for some taxa. In recognition of this reality, a second, more methodological goal of this work is to evaluate a series of strategies aimed at recovering accurate phylogenetic signal from potentially highly homoplasious mtDNA data. We are particularly interested in the case when increasing phylogenetic resolution through dense taxonomic sampling (Hillis, 1996; Pollock et al., 2002; Zwickl and Hillis, 2002) is impossible. Although our taxonomic sampling of the Trionychidae is complete, most of the modern softshell turtle taxa are representatives of old and very divergent lineages (Meylan, 1987). Thus, due to biological realities, our sampling consists of one or a few remaining repre-

sentatives of several ancient lineages across a relatively deep phylogenetic tree. In the case of softshells, additional sampling simply is not possible, because the set of surviving species is itself a sparse tree. Using our combined data set as a strong estimate of the best tree, we retrospectively evaluate several strategies to ask whether any one approach outperforms others for extracting phylogenetic signal from highly homoplasious mtDNA data.

#### *Previous Phylogenetic Hypotheses and Taxonomy*

The softshell turtles (*Trionychidae*) are an ancient, morphologically bizarre, and geographically widespread group of turtles characterized by reduction of the bony elements of the shell and complete loss of the keratinized, carapacial scutes that are characteristic of most other turtles. They include some of the largest (over 100 kg; Pritchard, 2001) and most endangered (Van Dijk et al., 2000) turtles in the world. Extant *trionychids* occur in North America, Europe, Africa, Asia, and the East Indies (Iverson, 1992). Fossil forms are also known from Australia (Gaffney, 1979a). The fossil record of *trionychids* is extensive (Romer, 1968), with some fossil taxa from as early as the late Cretaceous (Kordikova, 1991; Chkhikvadze, 2000) classified within modern genera. This fossil record and the highly autapomorphic morphologies of extant taxa (Meylan, 1987) suggest that the crown group may be evolutionarily ancient. The monophyly of the *Trionychidae* has never been questioned; however, the relationship of softshell turtles to other turtles has been controversial (reviewed by Gaffney and Meylan, 1988; Shaffer et al., 1997; Fujita et al., 2004). Recent molecular studies (Shaffer et al., 1997; Starkey, 1997; Fujita et al., 2004; Krenz et al., unpublished results) strongly support a sister relationship between the Austral/New Guinea pig-nosed turtle, (*Carettochelidae: Carettochelys insculpta*) and the softshell turtles, and place this clade (*Trionychoidea* of Shaffer et al., 1997; converted clade name *Trionychia* of Joyce et al., 2004) as the sister group of all other living *cryptodires*, and possibly to all other living turtles (Krenz et al., unpublished results). Based on fossil evidence and inferences from molecular data, the split between *Trionychia* and all other turtles is estimated to have taken place approximately 90 to 120 million years ago (Shaffer et al., 1997).

Our current conception of the relationships within the *Trionychidae* is based on Meylan's (1987) analysis of morphological characters of the skull, shell, and postcranial skeleton (see Fig. 1). Flap-shelled turtles, which can hide their feet under flaps of skin projecting from the plastron, have long been considered unique among softshell turtles (Boulenger, 1889; Lydekker, 1889) and are referred to the subfamily *Cyclanorbinae* (Meylan, 1987). All other softshell turtles are classified in the subfamily *Trionychinae* (Hummel, 1929). The monophyly of the flap-shells has been questioned (de Broin, 1977); however, Meylan (1987) describes 12 shared derived morphological characters for *Cyclanorbinae* and an additional 9 for *Trionychinae*, strongly supporting the reciprocal monophyly of the two subfamilies. Within *Cyclanorbinae*, Meylan

recognized four species in two African genera, *Cyclanorbis* and *Cycloderma*, which he placed in the tribe Cyclanorbini and one species in the genus, *Lissemys*, endemic to the Indian subcontinent for which he erected the tribe Lissemydini. The monophyly of each of these two groups has never been questioned.

In contrast, the taxonomy and phylogenetic relationships within *Trionychinae* have been far more controversial. Until Meylan (1987), all *trionychine* softshell turtles with the exception of the Southeast Asian giant genera (*Chitra* and *Pelochelys*) were included in a single “wastebasket” genus *Trionyx*. No evidence for the monophyly of *Trionyx* had ever been assembled and Gaffney (1979b) asserted that the all-inclusive genus *Trionyx* was based on plesiomorphic characters and that the continued use of “*Trionyx*” for all non-(*Chitra*, *Pelochelys*) trionychines was equivalent to “*Trionychidae* sp. eq.”. Meylan reclassified the 15 species formerly comprising *Trionyx* into nine genera, with the goal of a purely cladistic classification. To accomplish this goal, he resurrected seven genera, erected one novel genus, and left *Trionyx* as a monotypic genus containing *Trionyx triunguis*. This increased the total number of trionychine genera from 3 to 11, 8 of which were considered to be monotypic. Meylan grouped these 11 genera into four tribes: Chitrini, *Aspideretini*, Trionychini, and Pelodiscini (Fig. 1). Although Meylan considered the content and monophyly of each of these four clades to be well established, he was not able to resolve the relationships among the tribes. Meylan’s novel classification has been widely accepted, although the single large genus *Trionyx* is sometimes still used for most *Trionychine* softshell turtles (e.g., Nie et al., 2001; Plummer, 2001).

## MATERIALS AND METHODS

### Taxonomic Sampling and Laboratory Protocols

Based on previous studies of turtle phylogeny (Gaffney and Meylan, 1988; Shaffer et al., 1997; Starkey, 1997; Fujita et al., 2004; Krenz et al., unpublished results), *Carettochelys insculpta* was chosen as the most appropriate outgroup to the softshell turtles. We generated molecular data for *C. insculpta* and for 23 of 26 recognized species of softshell turtles, including representatives of all recognized genera. Tissues were not available for *Rafetus swinhoei*, *Aspideretes nigricans*, and *Aspideretes leithii*, all of which are critically endangered and extremely rare (Van Dijk et al., 2000). We also reanalyzed the morphological data from Meylan’s (1987) monographic treatment of the family. Details of morphological characters and specimens examined are provided in Meylan (1987). Meylan’s study includes morphological data for the three species for which we lack molecular data, but lacks morphological data from species which Meylan viewed as conspecific (*Lissemys scutata* and *L. punctata* [Webb, 1982]) or which have been recognized since Meylan’s work (*Chitra chitra* [Nutaphand, 1986; McCord and Pritchard, 2002], *Chitra vandijki* [Engstrom et al., 2002; McCord and Pritchard, 2002], and *Pelochelys bibroni* [Webb, 1995]). By combining these two data sets, we have complete taxo-

nomic sampling at the species level, with a full data set for 19 species, only morphological data for 3 species, and only molecular data for the remaining 4.

Samples for the following taxa were obtained from live animals in the private collection of William P. McCord: *Amyda cartilaginea* (Thailand), *Aspideretes hurum* (Dacca Market, Bangladesh), *Aspideretes gangeticus* (Dacca Market, Bangladesh), *Carettochelys insculpta* (South coast of Irian Jaya, Papua, Indonesia), *Chitra chitra* (Thailand), *Chitra indica* (Bangladesh), *Chitra vandijki* (Riuli Market, Yunnan Province, China; animal collected in Myanmar), *Cyclanorbis elegans* (Benin), *Cyclanorbis senegalensis* (Togo), *Cycloderma aubryi* (Gabon), *Cycloderma frenatum* (Lake Malawi), *Dogania subplana* (Panang, Malaysia), *Lissemys punctata* (India), *Lissemys scutata* (Myanmar), *Nilssonina formosa* (Myanmar), *Palea steindachneri* (China-Vietnam border), *Pelochelys bibroni* (South coast of Irian Jaya, Papua, Indonesia), *Pelochelys cantorii* (“Thailand,” either Menona, Cambodia, or Peninsular Thailand), *Pelodiscus sinensis* (Shanghai, China), and *Trionyx triunguis* (Liberia). Blood samples from *Apalone ferox* (Palm Beach County, Florida), *Apalone mutica* (Escambia River just north of State Road 4, Escambia County, Florida), and *Apalone spinifera aspera* (Ochlocknee River, Whitehead Landing, Liberty County, Florida) were collected by Paul Moler as part of long-term mark recapture studies. *Apalone spinifera emoryi* (introduced to the University of California Davis Arboretum Waterway, Yolo County, California [Spinks et al., 2003]), and *Rafetus euphraticus* (CAS 228508, Euphrates River, Biricik, Turkey) were field collected by TNE.

Blood and tissue samples were stored at 4°C in lysis buffer (White and Densmore, 1992). Genomic DNA was extracted by standard phenol/chloroform techniques (Palumbi, 1996) and stored at -20°C. Polymerase chain reaction (PCR) was conducted in 15 or 25 µL volumes containing 0.5 mM of each primer, 0.125 mM of each dNTP, 0.25 mM MgCl<sub>2</sub>, 0.5 M betaine, and 0.5 to 0.75 U *Taq* DNA polymerase using primers described in Table 1. Thermal cycle profile consisted of a 3-min initial denaturation at 94°C followed by 35 cycles of denaturation for 30 s at 94°C, annealing for 45 s at 50°C (mtDNA) or 60°C (nDNA intron), extension for 1 min at 72°C with a final 3-min extension at 72°C following the last cycle. Negative controls were used in all amplifications to check for possible contamination. Unincorporated primers and dNTPs were removed either using Millipore Ultrafree MC 30,000 NMWL filters or enzymatically using exonuclease 1, shrimp alkaline phosphatase treatment (Amersham Pharmacia Biotech), and sequenced at the U.C. Davis Division of Biological Sciences DNA Sequencing Facility (<http://dnaseq.ucdavis.edu/>) with an ABI 377 or ABI 3100 automated sequencer. All DNA sequences were confirmed either by sequencing both the forward and reverse strands of a single PCR product or by sequencing the forward strand of two PCR products from two different reactions from the same individual. Purported *cytb* sequences from *Rafetus euphraticus* and the three species of *Apalone* each had a one base pair indel

immediately prior to the stop codon suggesting that these sequences may have come from nuclear pseudogenes. To confirm that these sequences are mitochondrial cytochrome *b*, we sequenced the 3' end of *cytb* from multiple PCR products including long PCR products (~2300 bp) spanning the adjacent ND6 and control region from several DNA extractions from each of the four individuals and from three different individuals of *R. euphraticus*. We also compared patterns of sequence evolution in the purported *cytb* sequence with known *cytb* pseudogenes. The presence of a single genomic copy of the R35 gene was confirmed by Southern blot analysis of genomic DNA (Fujita et al., 2004). mtDNA sequences were aligned by eye using SeqEd V.1.0.3 (Applied Biosystems), and R35 intron sequences were aligned using clustal X (Thompson et al., 1997). Sequences were deposited in GenBank (accession numbers: *cytb*, AY259546–AY259570, ND4 AY259596–AY259615, R35 intron AY259571–AY259595). Aligned sequence data are available on treebase (<http://www.treebase.org>).

#### *Tests for Excessive Homoplasy ("Saturation")*

Molecular data were tested for substitutional saturation by plotting observed pairwise distance for transitions and transversions for each pair of taxa against the corrected distance estimated from maximum likelihood. Unsaturated data are expected to increase linearly, whereas saturated data are expected to show a distinct plateau at higher levels of divergence (Irwin et al., 1991; Graybeal, 1994). To construct a reproducible criterion for "saturation," we fitted a 2nd order polynomial regression line to the saturation plots. If the slope of this regression line was zero or negative for comparisons within the ingroup taxa, we considered the data saturated. For the mitochondrial protein-coding genes, saturation plots were constructed for each codon position separately, and *cytb* was further divided into structural partitions of intermembrane, transmembrane, and matrix (Degli et al., 1993; Griffiths, 1997). This resulted in two saturation plots for the intron (ti and tv), six from ND4 (ti and tv for each of 3 codon positions), and 18 from *cytb* (ti and tv for 3 codon positions in 3 structural regions). Data in which transitions but not transversions were considered saturated were excluded by recoding nucleotides as purine or pyrimidine and appending this recoded data to the end of the data matrix. Analyzing recoded data rather than using transition matrixes to analyze the original data allows parsimony searches to proceed more rapidly and also allows analysis of transitionless data using likelihood criteria whereas the use of transition matrixes does not.

#### *Tests for Base Compositional Bias*

We tested the possibility that our phylogenetic analyses were misled by base composition bias using the base stationarity test implemented in PAUP\* version 4b10 (Swofford, 2002). Tests were carried out using only variable sites for each gene individually, for combined mtDNA, and for all molecular data. Although the test

does not constitute a rigorous test of base composition bias because it ignores correlation of characters due to phylogenetic structure and lacks power, the qualitative assessment of the degree and direction of differences in base frequencies among taxa is still informative. In cases when a strong difference in base composition among taxa was detected, the direction of bias was compared to phylogeny inferred using character-based methods to determine if conflicts in our analyses reflected greater similarity in base composition rather than phylogenetic history.

#### *Phylogenetic Analyses*

Phylogenetic analyses using maximum parsimony and maximum likelihood were performed using PAUP\* version 4b10 (Swofford, 2002) with heuristic searches using TBR branch swapping. Support for nodes was assessed using nonparametric bootstrap analysis based on 1000 pseudoreplicates with 10 random sequence additions for all parsimony analyses and on 100 pseudoreplicates with 1 random sequence addition for all likelihood analyses. Because a goal of our analysis was to evaluate the ability of different analytic methods to recover phylogenetic information from both saturated and unsaturated data partitions, we analyzed data from each gene separately and combined under a variety of conditions.

Maximum parsimony analyses were performed for the morphological, *cytb*, ND4 (including 23 bp of tRNA<sup>His</sup>), combined mtDNA, combined mtDNA and morphology, intron, combined mitochondrial and intron, and combined molecular and morphological data using several weighting schemes. In equally weighted analysis all characters and all types of character changes were assigned a weight of 1. We also conducted analyses using step matrices to exclude transitions entirely, or to differentially weight transitions and transversions. We chose ti/tv weights by using maximum likelihood to estimate a transition bias ( $\kappa$ ) for all data and for each codon position within each gene separately on the most likely tree (Voelker and Edwards, 1998). Step matrices were used to give transitions a weight equal to  $1/\kappa$  and transversions were given a weight of 1. In "weeded" analysis, saturation plots were used to identify and exclude a priori any data suspected to be highly homoplasious and thus potentially misleading.

Maximum-likelihood analyses were performed on *cytb*, ND4 (including 23 bp of tRNA<sup>His</sup>), combined mtDNA, intron, and combined mitochondrial and nuclear data sets. Additional analyses were carried out for "weeded" mtDNA data sets from which suspected saturated transitions were excluded by recoding sites as purine-pyrimidine. For each data set, initial model choice and parameter values were estimated using Modeltest Version 3.06 (Posada, 2001). These were used to construct an initial tree, which was then used to estimate new parameter values. This process of iterative parameter estimation was repeated until two iterations returned the same parameter values. The parameters from this iteration were then used in all further analyses. Aligned

sequence data and details of models used for each partition are in a nexus file available from the senior author or at <http://www.treebase.org>.

Because we detected significant base composition bias, and significant among site rate variation, we also performed LogDet paralinear distance analysis, which is less likely to be misled by nonstationary base composition (Gu and Li, 1998). As with maximum-likelihood and Bayesian analyses, LogDet+I analyses were performed on *cytb*, ND4 (including 23 bp of tRNA<sup>his</sup>), combined mtDNA, intron, and combined mitochondrial and nuclear data sets. The proportion of invariant sites for each analyses was estimated separately for each gene or combination of genes using GTR+I model of sequence evolution. Estimated values for I were *cytb* (0.48772), ND4 (0.42476), mtDNA (0.46242), intron (0.244705), combined mtDNA and nuclear (0.51645).

Bayesian analyses were performed using Mr. Bayes V. 3.0 (Huelsenbeck and Ronquist, 2001). Analyses were performed on separate *cytb*, ND4 (including 23 bp of tRNA<sup>his</sup>), combined mtDNA, intron, combined mitochondrial and nuclear, and combined molecular and morphological datasets. In combined analyses, data were partitioned by gene and by codon position within protein coding genes. Molecular partitions were analyzed with a GTR+I+G model of sequence evolution with parameters for each partition estimated separately for each molecular data partition, and morphological partition was analyzed with Lewis' (2001) maximum-likelihood approach to modeling discrete morphological character data. Default priors were used in each analysis using four, heated MCMC chains. We started each analysis from two different random starting points to confirm convergence and mixing and ran each analysis 4,000,000 generations, saving trees every 100 generations (40,000 saved trees total). The first 1,000,000 generations (10,000 trees) were discarded as "burn in," and the remaining 30,000 sampled generations were used to estimate posterior probabilities of tree topology and parameters values.

We used partitioned maximum likelihood (DeBry, 1999; Wilgenbusch and de Queiroz, 2000; Caterino et al., 2001) to evaluate support for various phylogenetic hypotheses under the best possible model of molecular evolution, because it is possible for maximum-likelihood analyses to favor incorrect topology if the model of molecular evolution used is not correct (Buckley and Cunningham, 2002). This could occur if the global likelihood model selected does not adequately describe the heterogeneous evolutionary processes of different partitions of the data, and we did not want our acceptance or rejection of a particular hypothesis to be compromised by poor model choice. We also used partitioned maximum-likelihood to identify which partitions supported conflicting topologies. For partitioned likelihood analyses, data were divided into eight partitions consisting of (1) nuclear intron, (2) *cytb* 1st position, (3) *cytb* 2nd position, (4) *cytb* 3rd position, (5) ND4 1st position, (6) ND4 2nd position, (7) ND4 3rd position, and (8) tRNA<sup>his</sup>. We used Modeltest v. 3.06 (Posada, 2001) to select the best model of sequence evolution

and parameter values for each data partition. Details of models used for each partition are available from the senior author or in an aligned data file on treebase (<http://www.treebase.org>). We calculated partitioned likelihood scores for each phylogenetic hypothesis by estimating likelihood scores for each partition separately and then summing across all partitions. Hypotheses tested using partitioned analysis included (1) the best trees from our maximum-likelihood, maximum-parsimony, weeded maximum-parsimony, and Bayesian analyses; (2) Meylan's (1987) preferred topology; (3) trees representing each of Meylan's four trionychine tribes, Chittrini, *Aspideretini*, Trionychini, and Pelodiscini; and (4) the most likely tree containing a monophyletic "*Trionyx*" in the pre-Meylan sense of that name. The most likely tree containing each of these nodes of interest were chosen by constraining the monophyly of each of these groups in a maximum-likelihood search with the combined molecular data.

Alternative topologies were tested using parametric bootstrapping procedures outlined by Huelsenbeck et al. (1996). Test trees and simulation model parameters were selected using PAUP\* to estimate parameters for GTR+I+G model of sequence evolution in maximum-likelihood searches with topological constraints consistent with each hypothesis. These trees and model parameters were then used as to simulate 1000 data matrices equal in size to the original matrix using the Genesis module in Mesquite v. 0.996 (Maddison and Maddison, 2003). PAUP\* was then used to conduct two parsimony searches for each simulated data matrix, either constrained to the hypothesis being tested or unconstrained. Differences in tree length for constrained and unconstrained searches for each of the 1000 simulated matrices were calculated and plotted as histograms using Mesquite v. 0.996 (Maddison and Maddison, 2003). This serves to build a null distribution of tree length differences between two potential topologies. If the difference between constrained and unconstrained topologies in the original data set falls outside the 95% confidence interval of this distribution, then the hypothesis that the constraint tree constitutes the true evolutionary history is rejected in favor of the shorter unconstrained topology.

## RESULTS

The primers described in Table 1 consistently amplified single gene products of appropriate size from all softshell turtles. With one exception, all mitochondrial protein-coding sequences were successfully translated into proteins similar to published turtle sequences (Zardoya and Meyer, 1998; Kumazawa and Nishida, 1999; Mindell et al., 1999). A single-base pair indel at the 3' terminus of *cytb* was detected in both forward and reverse sequencing reactions of all PCR and long PCR products in all members of the *Apalonina* clade. These purported *cytb* sequences do not show signature patterns of sequence evolution common to nuclear pseudogenes, including a 5% to 10% decrease in rate of

TABLE 1. PCR primers used in this study. Position is in reference to the complete mitochondrial sequence of *Dogania subplana* (Farajallah et al., AF366350).

Primer name	Position	Sequence	Reference
ND4 672(f)	10904	TGACTACCAAAAGCTCATGTAGAAGC	Engstrom et al., 2002
Hist (r)	11628	CCTATTTTITAGAGCCACAGTCTAATG	Aravelo et al., 1994
Gludg (f)	14165	TGACTTGAARAACCAAYCGTTG	Palumbi, 1996
CB2 (r)	14591	CCCTCAGAATGATATTTGTCCTCA	Palumbi, 1996
CB94lt (f)	14487	TGCATCTACC TTCACATYGG MCG	Shaffer et al., 1997
CB3 (r)	14999	GGCAAATAGGAAATATCATTC	Palumbi, 1996
CB534(f)	14718	GACAATGCAACCCTAACACG	This study
CB649(r)	14834	GGGTGGAATGGGATTTTGTGTC	This study
CB791(f)	14976	CACCMGCYAAACCCACTATC	This study
Tcytbthr(r)	15355	TTCTTTGGTTTACAAGACC	This study
ND6 346F	13938	GAATAAGCAAAAACCACTAACATACCCCC	This study
TCR500	16616	CCCTGAAGAAAAGAACCGAGGCC	This study
R35Ex1 (f)	R35Exon1	ACGATTCTCGCTGATTCTTGC	Fujita et al., 2004
R35Ex2 (r)	R35Exon2	GCAGAAAACCTGAATGTCTCAAAGG	Fujita et al., 2004

divergence (e.g., Arcander 1995; DeWoody et al., 1999; Lü et al., 2002), high incidence of indels resulting in multiple-phase shift or stop codon mutations (Bensasson et al., 2000), decrease in ti/tv ratio from typically high mtDNA ratio to ~2:1 (DeWoody et al., 1999), and loss of differences in substitution pattern among (former) codon positions (Bensasson et al., 2000). The sequences also show a paucity of guanine, which is typical of mtDNA protein-coding genes and has been used as a criterion for identifying authentic mtDNA (Macey et al., 1997a, 1997b; Schulte et al., 2003). Extra untranslated nucleotides have been described in the mtDNA ND3 gene in some birds and a turtle (Mindell et al., 1998). Given the weight of this evidence, we conclude that these sequences are authentic mitochondrial *cytb*. The R35 intron sequences ranged from 975 to 1034 bp. However, indels ranging from 1 to 24 bp were common, yielding an aligned sequence matrix of 1063 nucleotide positions.

#### Reanalysis of Morphological Data

Meylan's (1987) study predated the intensive searching of tree space and statistical analyses that are now standard in phylogenetics. We provide both bootstrap (BP) and decay values, based on Meylan's dataset, in Figure 1. Our reanalysis of Meylan's data produced six equally most parsimonious trees of length 193 (CI 0.446, RI 0.621, RC 0.277, HI 0.554). The topology of the bootstrap consensus tree is identical to the strict consensus of these six trees (not shown). This consensus tree is largely consistent with the preferred tree upon which Meylan based his taxonomy. In Figure 1 we show Meylan's (1987) preferred tree, with BP and decay values from our consensus tree. There is strong support for the monophyly of both *Cyclanorbinae* and *Trionychinae*, the sister relationship of *Chitra* and *Pelochelys* (BP = 99), the monophyly of *Rafetus* (BP = 89), and for the (*Apalone spinifera*, *A. mutica*) clade (BP = 77). However, of the five nonmonotypic genera recognized, only *Rafetus* has a reasonable level of bootstrap support, with others ranging from 0% to 47%. Two of the five nonmonotypic tribes proposed by Meylan are weakly recovered as nonmonophyletic (indicated in Figure 1 by negative Bremer support at those nodes),

with *Nilssonina* falling outside the *Aspideretini*, and the *Cyclanorbini* recovered as paraphyletic with respect to *Lissemys*. The greatest bootstrap support for any of the three monophyletic tribes is 46% (*Chitridini*). As noted by Meylan, for the relationship among the four, trionychine tribes are poorly resolved. Overall, our reanalysis of Meylan's morphological data set provides weak support for some of the named taxa in Meylan (1987), and weak conflict with others.

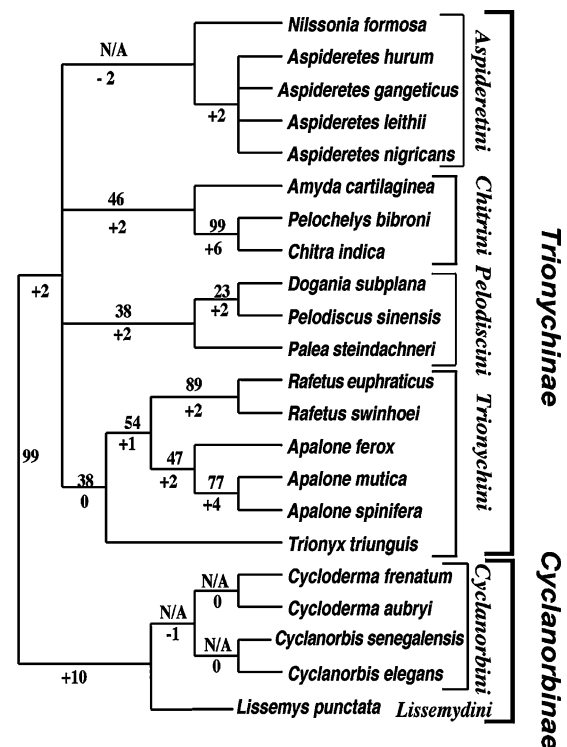


FIGURE 1. Meylan's (1987) preferred tree upon which he based his phylogenetic classification of the softshell turtles. Bootstrap support and decay indices (Bremer support) from our reanalyses are shown above the node and below the node respectively. Negative decay index indicates that a node was not found in the most parsimonious tree.

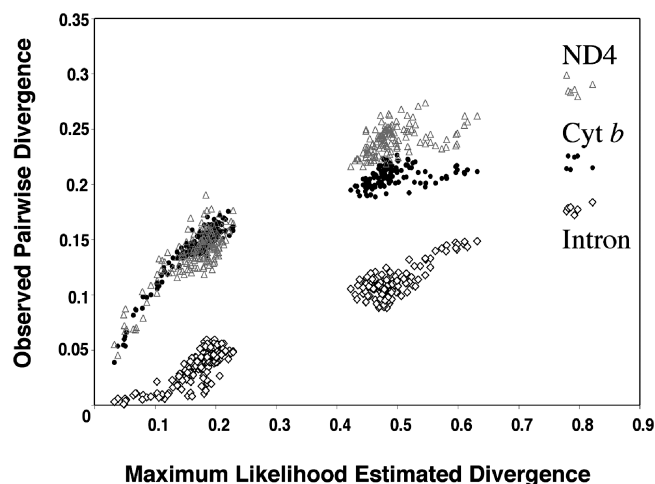


FIGURE 2. Saturation plot showing the pairwise divergence in the ND4, *cytb*, and R35 introns. The x-axis represents pairwise divergence estimated by maximum likelihood. The y-axis is the observed divergence. Comparisons for ND4 are shown with open triangles, *cytb* with closed circles, and the R35 intron in open squares.

#### Molecular Data

**Tests for saturation.**—Saturation plots (Figs. 2 and 3) show that the two mitochondrial genes evolve at a substantially faster rate than the nuclear intron. Transitions at the third positions of all process partitions of both

mtDNA genes conformed to our saturation criteria (only *cytb* shown in Fig. 3). Transitions at 1st positions in both the matrix and transmembrane, but not in the intermembrane, portions of the *cytb* data were saturated. Based on this evidence of substitutional saturation, these partitions were excluded from some maximum-parsimony and maximum-likelihood analyses. The three process partitions in the *cytb* data showed evidence of heterogeneous patterns of substitution (Fig. 3). Overall, intermembrane sites showed the slowest substitution rates in 1st and 2nd positions, and transmembrane sites showed the fastest rates.

**Tests for base compositional bias.**—Base stationarity tests showed significant differences among taxa in base composition bias in the mtDNA data (*cytb*:  $\chi^2 = 127.66$  [ $df = 72$ ],  $P = 0.000059$ ; ND4:  $\chi^2 = 118.43$  [ $df = 72$ ],  $P = 0.00048$ ; combined mtDNA:  $\chi^2 = 214.69$  [ $df = 72$ ],  $P = 0.00$ ), but not in the intron ( $\chi^2 = 19.75$  [ $df = 72$ ],  $P = 1.00$ ) or combined data ( $\chi^2 = 62.27$  [ $df = 72$ ],  $P = 0.79$ ). This base composition bias in the mtDNA cannot explain the topological conflict between our mtDNA and intron data. The case in which conflict between mitochondrial and nuclear genes is most apparent regards the placement of *Rafetus euphraticus* relative to the Asian clade (*Amydona*) and the North American softshells (*Apalone*). Phylogenetic analysis of mtDNA places *Rafetus euphraticus* as sister to the *Amydona*; however, the mtDNA base composition of *Rafetus* is more similar to that of the three

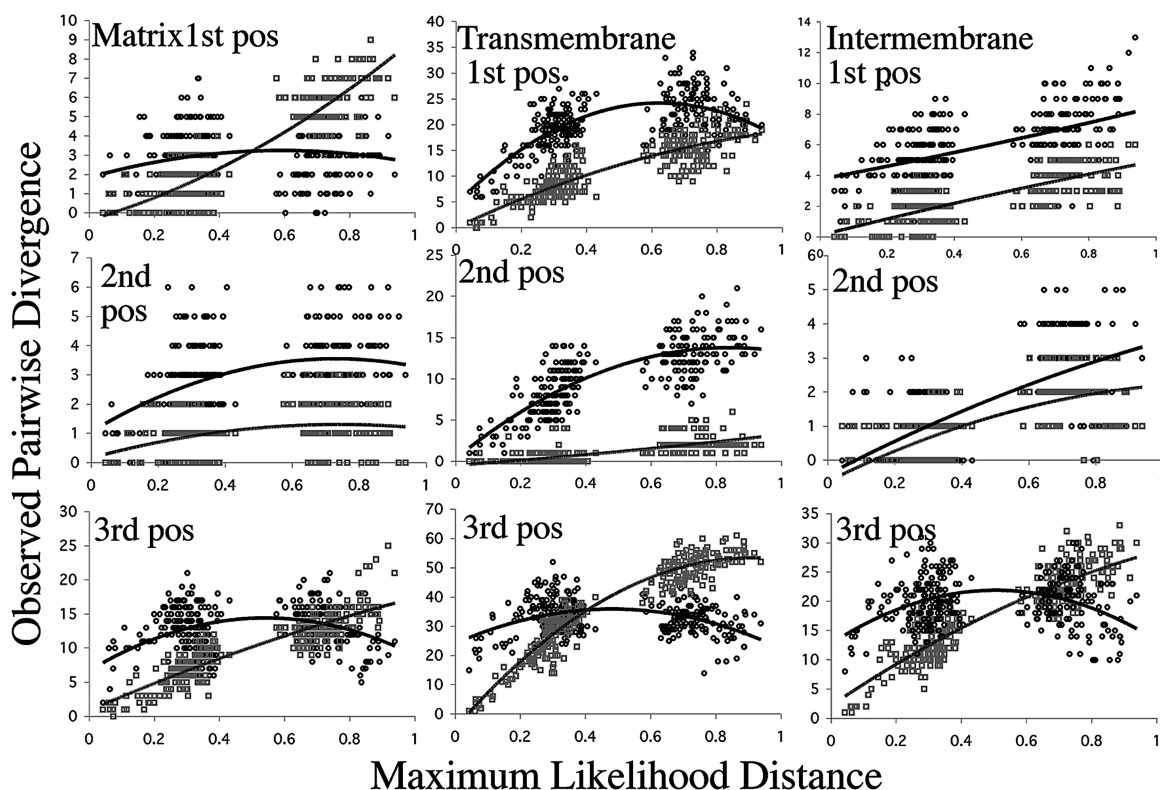


FIGURE 3. Saturation analyses of transitions and transversions at 1st, 2nd, and 3rd codon positions in the intermembrane, transmembrane, and matrix partitions of the mitochondrial *cytb* gene. The x-axis represents pairwise distance estimated by maximum likelihood, the y-axis is the number of observed substitutions. Transitions are shown as dark open circles transversions are open gray squares. The trend line is a best-fit 2nd degree polynomial.

species of *Apalone*. Similarly, phylogenetic analysis of ND4 data renders *Cyclanorbini* paraphyletic by placing *Cycloderma* and *Lissemys* as sister taxa to the exclusion of *Cyclanorbis*; however, the base composition of *Cycloderma* is closer to that of *Cyclanorbis* than to *Lissemys*. These results are consistent with a recent simulation study showing that the level of base composition bias needed to mislead phylogenetic methods in simulated data sets is far higher than that normally found in nature (Conant and Lewis, 2001) and much higher than in our data.

**mtDNA.**—In parsimony, likelihood, LogDet, and Bayesian analyses, ND4 provided strong support (BP > 90, PP > 95) for several tip nodes, but very low support (<50) for most deep nodes (Fig. 4, see Appendix 1 for detail, available at the Society of Systematic Biologists Website, <http://systematicbiology.org>). Equally weighted parsimony analysis of the *cytb* data (Appendix 1) also left the deeper nodes within the trionychines completely unresolved. However, maximum-likelihood (Appendix 1) and Bayesian (Fig. 4, Appendix 1) analyses of *cytb* did recover some strongly supported deep nodes. The two mitochondrial genes provide strong, concordant support for some nodes at the tips of the tree. However, the two mitochondrial genes, which presumably share a single genealogy, provide conflicting support for opposite relationships among the three species of North American softshell turtles (*Apalone ferox*, *A. mutica*, and *A. spinifera*). ND4 supports a sister relationship of *Apalone spinifera* and *A. mutica* (bootstrap support of 92 MP, 75 ML, 100 Bayes), whereas *cytb* supports the sister relationship of *A. spinifera* and *A. ferox* (80, 74, 77).

Analysis of the combined mtDNA places African cyclanorbines as monophyletic, *Trionyx triunguis* as sister to the Southeast Asian giants *Pelochelys* and *Chitra* (68, 84, 100), and weakly places *Rafetus euphraticus* as sister to the Asian clade (<50 ML, 79 Bayes). Relationships among the three species of *Apalone* are not resolved, with likelihood weakly supporting an (*Apalone spinifera*, *A. ferox*) clade and parsimony and Bayesian analyses supporting (*A. spinifera*, *A. mutica*). Overall, the mtDNA strongly supports monophyly of the *Cyclanorbinae* and *Trionychinae*, and of three major clades within *Trionychinae*, but is not able to resolve the relationships among these clades, and is not able to place *Rafetus*.

**mtDNA versus nuclear intron.**—Saturation plots, low bootstrap values, and weak or conflicting resolution of some nodes all indicate that homoplasy may be an issue with the mtDNA data (Fig. 3, Appendix 1). In contrast the linear accumulation of substitutions in the intron (Fig. 2) is accompanied by the resolution of deep nodes with high levels of almost homoplasy-free character support, suggesting that the R35 intron may provide a more reliable estimation of deep nodes in the softshell phylogeny (Graybeal, 1994). The R35 intron provided remarkably good resolution for deep nodes but relatively little information regarding relationships at the tips of the tree. Intron data strongly support the monophyly of *Trionychinae* and *Cyclanorbinae* (100 MP, 100 ML, 100 Bayes), monophyly (100, 100, 100), and pectinate struc-

ture of the Asian clade (>95, >97, 100 for all nodes within the clade), *Trionyx triunguis* as the sister of the Southeast Asian giants (*Chitra*, *Pelochelys*) (81, 91, 100), and *Rafetus euphraticus* as sister (99, 100, 100) to a monophyletic North American *Apalone* clade (100, 100, 100), not as part of the Asian clade as suggested by mtDNA. Within *Cyclanorbinae*, the intron provides strong support for the monophyly of *Lissemys* (99, 100, 100) and moderate support for *Cycloderma* (78, 70, 93), but the monophyly of *Cyclanorbis* and of the two African genera *Cyclanorbis* and *Cycloderma* is equivocal. The intron supports the sister relationship between *Apalone spinifera* and *A. ferox* (supported by the *cytb*, conflicting with ND4 and morphological data). The topology supported by the intron (Fig. 4) differs from the mtDNA topology in several key aspects, most notably in the placement of *Rafetus euphraticus*. This relationship is supported by 13 intron characters (11 of which have a consistency index of 1) receives bootstrap support of 100 in both maximum-parsimony and maximum-likelihood analyses, and has a Bayesian posterior probability of 100%.

Maximum-likelihood and Bayesian analyses of the combined, nearly homoplasy-free intron with the more variable mtDNA data converged upon a single set of relationships with strong support for both deep nodes and tip nodes (Figs. 4, 5). The conflicts within mtDNA and between mtDNA and the intron are resolved with African cyclanorbines monophyletic (80, 67, 99), and *Rafetus euphraticus* sister to the North American *Apalone* (—, 83, 100) and *A. spinifera* is sister to *A. ferox*, although support for this relationship is weak (60, 69, 75). The currently recognized tribes, Chittrini, Pelodiscini, Trionychini, and “*Trionyx*” in the pre-Meylan sense, do not appear as monophyletic groups. Each of the five topologies in which one of these groups was constrained as monophyletic was statistically rejected using parametric bootstrap analyses of the combined molecular data ( $P \ll 0.01$ ).

#### Combined Molecular and Morphology

By combining all morphological and molecular data, we obtained a phylogenetic hypothesis for all extant softshell turtles. The topology and bootstrap support for nodes based on parsimony analysis of the morphological/molecular data for the complete taxon matrix using 1/ $\kappa$  ti/tv weighting for the mtDNA are shown as the top number above each node in Figure 5. Bayesian posterior probabilities for combined morphological and molecular data are shown as the bottom number below each node. Bootstrap support and posterior probabilities from maximum-likelihood and Bayesian analyses of the 24 taxa with molecular data available are shown on the same tree. Support for virtually all nodes is high, and the only conflict between analysis of the combined molecular and morphological data set, and of the molecular data alone, regards alternative relationships among the three North American softshell turtles. Both parsimony and Bayesian analyses group *Apalone spinifera* with *A. mutica* to the exclusion of *A. ferox*, whereas molecular



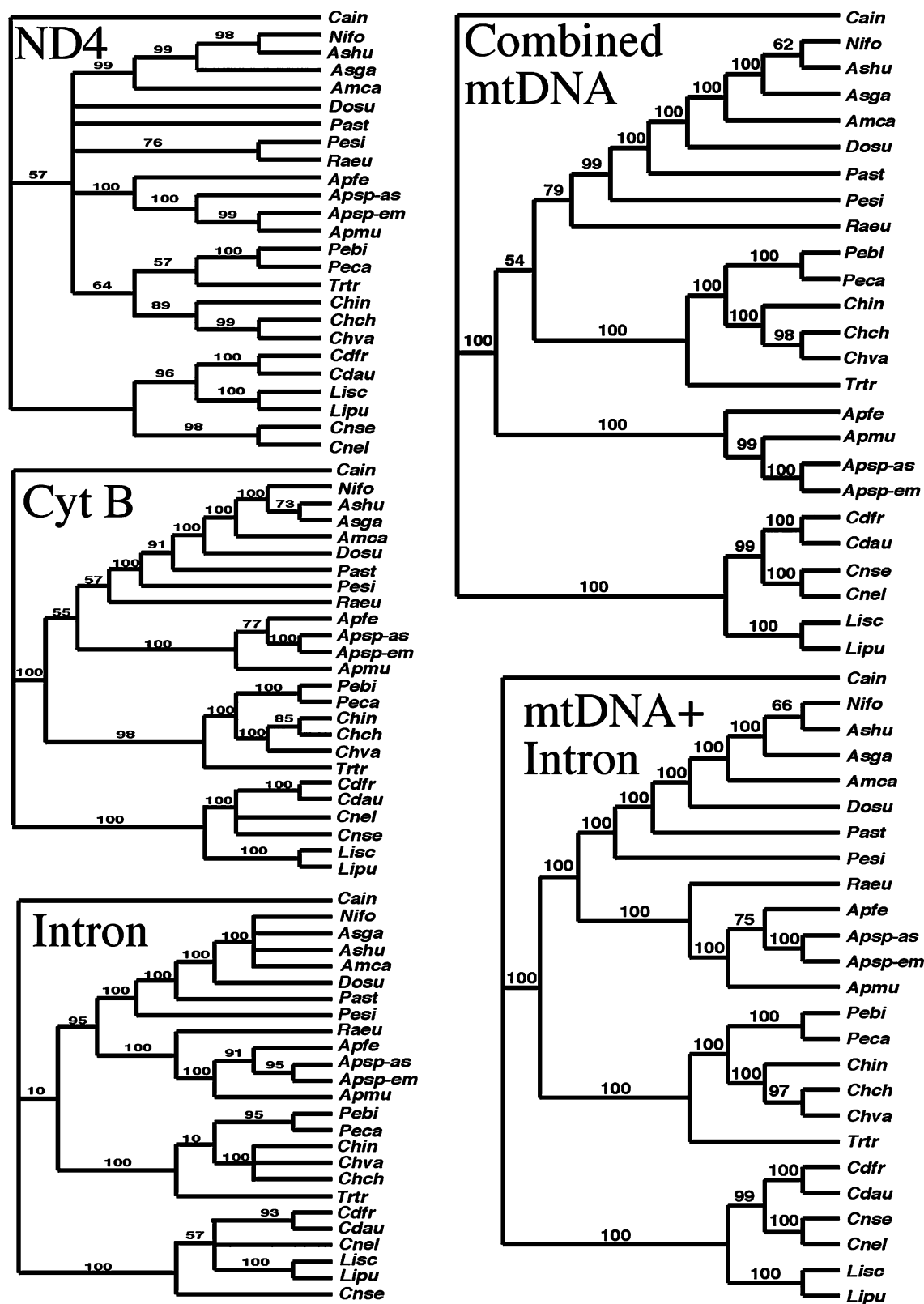


FIGURE 4. Phylogenetic trees for 23 softshell turtle species based on combined and separate analyses of mitochondrial and nuclear DNA data using Bayesian analyses. Numbers above the node are Bayesian posterior probabilities. Name abbreviations are the first two letters of the genus followed by first two letters of species name except for *Cyclanorbis* (Cn) and *Cycloderma* (Cd). Subspecies of *Apalone spinifera aspera* and *A. s. emoryi* are abbreviated *Apsp-em* *Apsp-as*.

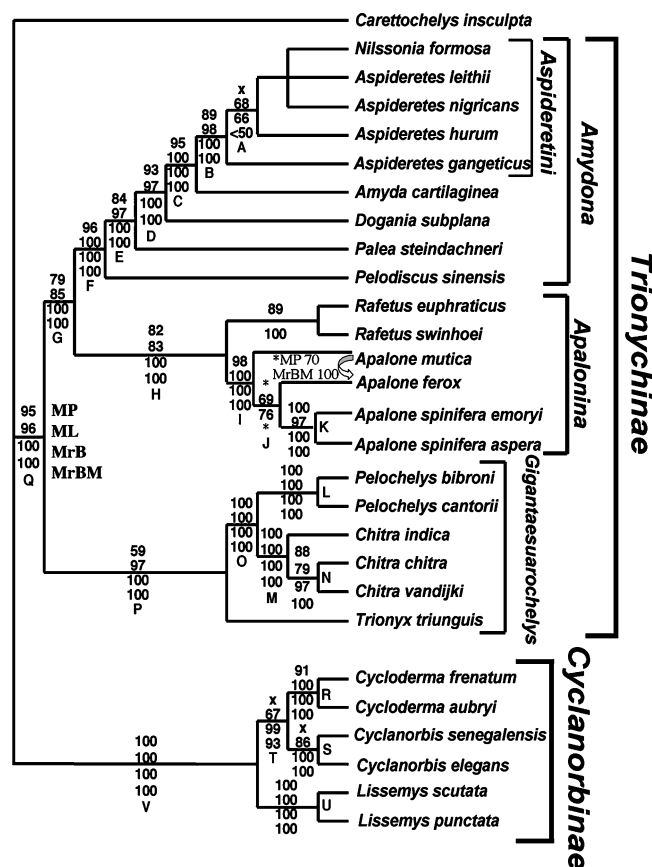


FIGURE 5. Our best estimate of the phylogenetic relationships of softshell turtles based on maximum-likelihood and Bayesian analyses of combined mitochondrial and nuclear DNA sequence data, 1/ $\kappa$ -weighted parsimony analyses of combined molecular and morphological data, and Bayesian analyses of combined molecular and morphological data. The bootstrap support from MP (all data) and ML (DNA data only) analyses are listed above the node, and the Bayesian posterior probability for DNA only (MrB) and for DNA and morphological data (MrBM) is below the node. The nodes lettered A–V are referred to in Appendix 1.

data weakly support the sister relationship of *A. ferox* and *A. spinifera* (shown in Fig. 5). Overall, we consider this a very strongly supported topology, and our best current estimate of the phylogeny of softshell turtles.

#### Getting the Most Out of Homoplasious Data: Weighting, Weeding, and Combining Data

Bootstrap support from various analyses of separate and combined mitochondrial data partitions and combined mitochondrial and nuclear partitions for each of the 22 nodes lettered A to V in the topology in Figure 5 are shown in Appendix 1. There are a few relationships, including the monophyly of *Trionychinae* and *Cyclanorbinae* (nodes Q and V), monophyly of *Apalone* (I), and sister relationship of *Pelochelys* and *Chitra* (N), that are well supported in all analyses of all partitions and combinations of data. Other deep nodes that are very strongly supported in the intron and combined analyses, such as placement of *Amyda*, *Dogania*, *Palea*, and *Pelodiscus* as suc-

cessive sister groups to (*Aspideretes*, *Nilssonina*) (nodes C, D, E, F), placement of *Rafetus* as sister to *Apalone* (H), and placement of *Trionyx* as sister to (*Chitra*, *Pelochelys*) (P), receive consistently weak support from equally weighted parsimony analysis of mitochondrial genes, both separately and when combined. In our evaluation of the efficacy of various analytical techniques, we assume that the combined tree presented in Figure 5 is "correct" and focus primarily on the ability of a given technique to recover these six difficult nodes (C, D, E, F, H, P). As a gross indicator of effectiveness, we summed bootstrap scores and calculated the average bootstrap score across the entire tree and for nodes C, D, E, F, H, P (bottom two rows of Appendix 1).

In 11 out of 12 cases, weighted and weeded parsimony analyses improved bootstrap support for the six difficult nodes when compared with equally weighted parsimony. Only transversion parsimony analysis of the combined mtDNA/nDNA data set did not increase overall bootstrap support for the key nodes. The effect of weighting schemes on bootstrap support across the entire tree was not as uniformly positive. Weighting or weeding increased overall tree bootstrap support in seven cases and decreased bootstrap support in five. Transversion parsimony was the least effective weighting scheme, resulting in a decrease in bootstrap support for the entire tree, in three of four data partitions and for key nodes in one of four partitions. In contrast, weeded parsimony and ti/tv weighting (both of which retain some information from transitions) increased support for the six difficult nodes in all partitions and increased support for the overall topology in all partitions except ND4. In similar weighted and weeded analyses of combined molecular and morphological data (not shown), all weighting schemes resulted in increased bootstrap support for the six key nodes relative to unweighted parsimony (average +149 BP points, +24.8 points per node). Weeding and ti/tv weighting also resulted in moderate increase across the entire tree (average +48 BP points, +2.2 points per node), but transversion parsimony resulted in a loss of support for the overall tree (–42 BP points, –1.9 points per node). Likelihood and Bayesian techniques recovered difficult nodes with levels of support that were much greater than equally weighted parsimony and usually greater than weeded and weighted parsimony. For all partitions except ND4, unweeded likelihood analysis returned the highest overall bootstrap support and either weeded or unweeded likelihood analysis returned the highest levels of support for the six key nodes.

It is important to note that although the sister relationship of *Rafetus* and *Apalone* (node H) is very strongly supported by intron data and by morphological data, no weeding/weighting scheme or model-based technique was able to recover the *Rafetus* and *Apalone* sister relationship using mitochondrial data alone. Only LogDet+I analyses successfully recovered this relationship from mtDNA data albeit very weakly (Appendix 1). Support for an alternative placement of *Rafetus* was generally weaker in weeded, weighted, and model-based analyses compared to equally weighted

analyses, indicating misplacement of *Rafetus* may be due to long-branch attraction, and that the problem is most severe under parsimony. In maximum-likelihood, LogDet, and Bayesian analyses of the combined mtDNA plus nDNA data set node H (*Rafetus*, *Apalone*) is recovered with high levels of support. This node is also recovered by ti/tv weighting and weeded parsimony analysis of combined molecular data (Appendix 1) and in Bayesian and parsimony analyses of combined morphological and molecular data (Fig. 5).

#### Partitioned Maximum Likelihood

Partitioned maximum-likelihood analysis resulted in a large overall improvement in likelihood scores compared with analysis using the global likelihood model (see Appendix 2 for detail, available at the Society of Systematic Biologists website, <http://systematicbiology.org>). There is no support from individual or summed partitions for the monophyly of the tribes Chitrini, Pelodiscini, or Trionychini (Fig. 1), monophyly of the previous concept of the genus "*Trionyx*" (all trionychines except for *Chitra* and *Pelochelys*), or the overall topology of Meylan's tree. Partitioned likelihood shows that support for the topology in which of *Rafetus* is not sister to *Apalone* comes exclusively from the mtDNA 3rd codon partitions, which are the most prone to long-branch attraction. There is no support for this topology from mtDNA 1st or 2nd codon positions or from the intron.

#### DISCUSSION

The two primary objectives of this study were to produce a well-supported phylogeny for all extant softshell turtles and to evaluate strategies for extracting phylogenetic signal from highly homoplasious mtDNA data. We first discuss key points of the phylogeny of softshell turtles and build a taxonomy with which that phylogeny can be effectively communicated. We then use this strongly supported phylogenetic taxonomy to discuss the analytical strategies of more general phylogenetic interest.

#### Phylogenetics and Taxonomy of Softshell Turtles

**Phylogeny.**—Softshell turtles are a morphologically unique, ancient, group of economically important turtles. They include the largest freshwater turtles in the world (Pritchard, 2001), and some of most threatened of any vertebrate species (Van Dijk et al., 2000). A strong phylogeny for the group is essential in assessing biodiversity, making sound management decisions, and understanding the evolution of their bizarre morphologies and biogeographic history. We have made major strides toward obtaining a complete phylogeny for the group. All forms of analysis, including maximum-likelihood and Bayesian analyses of molecular data and Bayesian and maximum-parsimony analyses of combined molecular and morphological data, converge on the well-supported set of relationships shown in Figure 5. We consider this to be our best operational hypothesis for relationships of softshell turtles. The only areas of un-

certainty within this set of relationships regard the potential paraphyly of *Aspideretes* with respect to *Nilssonina formosa*, the potential paraphyly of the African *Cyclanorbines* with respect to *Lissemys*, and the relationships among North American softshell turtles, *Apalone*.

Our analysis agrees with several key features of traditional (i.e., Meylan, 1987) ideas of softshell turtle systematics. We found unequivocal support for the monophyly of *Cyclanorbinae* and *Trionychinae*. Within *Cyclanorbinae*, we found strong support for the monophyly of *Cycloderma* and *Lissemys* and variable support for the monophyly of *Cyclanorbis*. Within *Trionychinae*, we found support for the long-held idea of a close relationship between the giant Southeast Asian genera *Chitra* and *Pelochelys* (Gray, 1873), for the monophyly of the North American softshell turtles (*Apalone*), and for a South Asian (*Nilssonina*, *Aspideretes*) clade (Meylan's *Aspideretini*). We also found strong support for the seemingly unlikely sister relationship of the Middle Eastern and East Asian *Rafetus* and the North American *Apalone*. Although the *Apalone*-*Rafetus* relationship was not recovered in analyses of mtDNA, it is strongly supported by the intron, weakly but consistently by the morphological data, and universally in all combined analyses.

Our analysis disagrees with other aspects of traditional softshell systematics. Using parametric bootstrapping, we are able to reject the monophyly of the currently recognized tribes Chitrini, Trionychini, and Pelodiscini, although for Chitrini and Trionychini, this is due to the misplacement of a single taxon from each. Another area in which our analyses conflict with traditional views of softshell turtle systematics, and show some internal conflict, is in the interrelationships among the three North American species of *Apalone*. One of the few statistically well-supported nodes in Meylan's morphological analysis is the sister relationship of the wide-ranging, broadly sympatric species *A. mutica* and *A. spinifera* to the exclusion of the allopatric, Florida endemic, *A. ferox*. Sequence data from ND4 also strongly support the sister relationship of *A. spinifera* and *A. mutica*. In contrast, a detailed phylogeographic study based on extensive sampling of all three species using *cytb* supports *A. spinifera* and *A. ferox* as sister taxa with strong (BP = 89) statistical support (Weisrock and Janzen, 2000). In our data, both *cytb* and the R35 intron support the Weisrock and Janzen (2000) topology with *A. spinifera* and *A. ferox* as sister taxa. For the time being, we provisionally favor the sister relationships of *Apalone spinifera* and *A. ferox* as our best hypothesis of the relationships of North American softshell turtles, based on support from global maximum-likelihood, Bayesian, and unweighted and unweeded parsimony analyses of our own molecular data and from the more densely sampled *cytb* phylogeny of Weisrock and Janzen (2000). We are currently attempting to resolve this conflict using increased taxonomic sampling and sequences from additional mtDNA genes.

**Taxonomic implications.**—The phylogeny proposed by Meylan (1987) has been the basis of softshell turtle taxonomy for the past 15 years. Although elements of this classification appear in our phylogeny (compare Figs. 1

and 5), few of Meylan's proposed tribes map to monophyletic groups. Given that Meylan's (and our) goal is a monophyly-based classification, we propose a novel phylogenetic classification of softshell turtles in accord with recommendations of the PhyloCode (de Queiroz and Cantino, 2001), following the protocol for conversion of rank-based names described by Joyce et al. (2004) (Table 2). As recommended by the PhyloCode, all clade names are italicized to clearly distinguish them from taxon names governed by the ICZN (1999) (e.g., *Trionychidae* versus *Trionychinae*). In assembling this classification, we have strived to simultaneously maintain nomen-

clatorial stability and accurately reflect our improved understanding of the phylogenetic relationships of these turtles. Phylogenetic definitions of converted clade names used in this classification are provided in Table 3.

The primary subdivision of *Trionychidae* into two clades, *Cyclanorbininae* and *Trionychinae*, is well supported in both Meylan's morphological data and our molecular data. Within *Cyclanorbininae*, we found some uncertainty regarding the monophyly of Meylan's tribe *Cyclanorbini* based on both molecular and Meylan's own morphological data. The clade *Cyclanorbini* is weakly supported and *Lissemys* contains a single genus and thus is redundant with *Lissemys*. We therefore recommend abandoning *Lissemys* but provisionally maintaining *Cyclanorbini* and *Cyclanorbis* as informal clade names pending additional investigation of the phylogenetic placement of *Cyclanorbis senegalensis*.

There is very little support for maintaining the current tribe designations within *Trionychinae*. Parametric bootstrapping of our molecular data rejects the monophyly of three of the four tribes proposed by Meylan (1987). Only *Aspideretini* is monophyletic in our analyses, and it is deeply nested within a group consisting of members of two of the other presently recognized tribes (Fig. 5). Several other monophyletic groups, which are ecologically, morphologically, or biogeographically coherent, and may therefore be worth recognizing taxonomically, do emerge from our analyses. The first of these is the clade of giant, often estuarine softshells containing *Trionyx*, *Pelochelys*, and *Chitra*. This group may be defined as consisting of all of the descendants of the most recent common ancestor of *Trionyx triunguis* and *Chitra indica*. All extant members are characterized by an extremely large adult body size, often exceeding 100 cm carapacial length (Pritchard, 2001) and frequent use of estuarine or marine habitats (Ernst and Barbour 1989). We suggest the clade name *Gigantaestuarocheles* for this group of giant, estuarine turtles (Table 2). Within *Gigantaestuarocheles*, the *Chitra*+*Pelochelys* clade has long been recognized as distinct. In a rank-free context Grey's (1873) subfamily name, *Chitraina* is available for this group, defined as all descendants of the most recent common ancestor of *Chitra indica* and *Pelochelys bibroni* (Table 3). *Gigantaestuarocheles* is sister to an unnamed group consisting of all of the descendants of the most recent common ancestor of *Amyda cartilaginea* and *Apalone ferox*. This large clade can be further divided into a diverse and pectinate clade of Asian/Indian softshells consisting of *Amyda cartilaginea*, *Dogania subplana*, *Palea steindachneri*, *Pelodiscus sinensis*, and *Aspideretini*. This clade is characterized by a small to medium size (Pritchard, 2001), and the presence of a strong symphyseal ridge located in a depression (Meylan 1987, character 95). *Amyda* is the oldest subordinate taxon name in this clade. To avoid confusion with the monotypic subtribe name *Amydina* erected by Meylan (1987), we suggest the name *Amydona* (signifying a larger clade) for the clade consisting of all of the descendants of the most recent common ancestor of *Amyda cartilaginea* and *Pelodiscus sinensis* (Tables 2, 3). *Amydona* is sister to a clade derived from the most recent common ancestor

TABLE 2. A rank-free phylogenetic taxonomy of softshell turtles based on the phylogenetic analyses presented in Figure 5. Definitions of converted clade names are given in Table 3. Definitions of new clade names are given in the text and in Table 3. Bibliographic documentation of original use of the name follows Meylan (1987) Table 22 except where indicated with an.\*

---

<i>Trionychidae</i> (Geoffroy-St.Hilaire, 1809a, 1809b) Bell, 1828
<i>Cyclanorbininae</i> (Hummel, 1929)
<i>Lissemys</i> (Malcom Smith, 1931)
<i>Lissemys punctata</i> (Lacépède, 1788)
<i>Lissemys scutata</i> (Peters, 1868)*
<i>Cyclanorbini</i>
<i>Cycloderma</i> (Peters, 1854)
<i>Cycloderma frenatum</i> (Peters, 1854)
<i>Cycloderma aubryi</i> (Dumeril, 1856)
<i>Cyclanorbis</i> (Gray, 1854)
<i>Cyclanorbis senegalensis</i> (Dumeril and Bibron, 1835)
<i>Cyclanorbis elegans</i> (Gray, 1869)
<i>Trionychinae</i> (Fitzinger, 1826) Lydekker, 1889
<i>Gigantaestuarocheles</i> (new clade name)
<i>Trionyx triunguis</i> (Forskål, 1775)
<i>Chitraina</i> (Gray, 1873)*
<i>Pelochelys</i> (Gray, 1864)
<i>Pelochelys bibroni</i> (Owen, 1853)*
<i>Pelochelys cantorii</i> (Gray, 1864)*
<i>Chitra</i> (Gray, 1844)
<i>Chitra indica</i> (Gray, 1831)
Unnamed Clade
<i>Chitra chitra</i> (Nutaphand, 1986)*
<i>Chitra vandijki</i> (McCord and Pritchard, 2002)*
Unnamed Clade
<i>Apalonina</i> (Meylan, 1987)*
<i>Rafetus</i> (Gray, 1864)
<i>Rafetus euphraticus</i> (Daudin, 1802)
<i>Rafetus swinhoei</i> (Gray, 1873)
<i>Apalone</i> (Rafinesque, 1832)
<i>Apalone mutica</i> (Le Sueur, 1827)
Unnamed Clade
<i>Apalone spinifera</i> (Le Sueur, 1827)
<i>Apalone ferox</i> (Schneider, 1783)
<i>Amydona</i> (new clade name)
<i>Pelodiscus sinensis</i> (Wiegman, 1835) Gray 1844
Unnamed Clade
<i>Palea steindachneri</i> (Siebenrock, 1906) Meylan, 1987
Unnamed Clade
<i>Dogania subplana</i> (Geoffroy-St.Hilaire, 1809a) Gray, 1844
Unnamed Clade
<i>Amyda cartilaginea</i> (Boddaert, 1770) Geoffroy-St.Hilaire, 1809a
<i>Aspideretini</i> (Hay, 1903) Meylan, 1987
<i>formosa</i> (Gray, 1869)
<i>gangeticus</i> (Cuvier, 1831)
<i>hurum</i> (Gray, 1831)
<i>leithii</i> (Gray, 1872)
<i>nigricans</i> (Anderson, 1875)

---

TABLE 3. Definitions of clade names used in the classification shown in Table 2. Following conventions established by de Quiroz and Cantino (2001) and Joyce et al. (2004), the most commonly used current names are assigned to the clade that most closely approximates its current meaning. The name of the relative stem clade is formed by the addition of the prefix "Pan." For example, "*Trionychidae*" as defined below includes all extant softshell turtles and any fossil forms arising from the last common ancestor of all extant softshell turtles. "*Pantrionychidae*" refers to this clade as well as all fossil forms, which share a more recent common ancestor with the crown clade than with its sister clade *Carettochelys insculpta*. Unless otherwise indicated, bibliographic documentation of original use of the name is provided in Meylan (1987), Table 22. Abbreviations: CCN, converted clade name; NCN, new clade name; ICN, informal clade name.

<i>Trionychidae</i>	As previously defined by Joyce et al. (2004), refers to the crown clade arising from the last common ancestor of <i>Trionyx</i> (orig. <i>Testudo</i> ) <i>triunguis</i> (Forskål, 1775), <i>Cyclanorbis senegalensis</i> (Dumeril and Bibron, 1835).
<i>Cyclanorbininae</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Cyclanorbis senegalensis</i> (Dumeril and Bibron, 1835) and <i>Lissemys punctata</i> (Lacépède, 1788).
<i>Cyclanorbini</i> (ICN)	Informal clade name for Meylan's (1987) tribe of the same name containing the taxa <i>Cycloderma frenatum</i> (Peters, 1854), <i>Cycloderma aubryi</i> (Dumeril, 1856), <i>Cyclanorbis senegalensis</i> (Dumeril and Bibron, 1835), and <i>Cyclanorbis elegans</i> (Gray, 1869). Not formally defined here because of ambiguity regarding the relationship of <i>Cyclanorbis senegalensis</i> (Dumeril and Bibron, 1835) to other members of <i>Cyclanorbininae</i> .
<i>Cyclanorbis</i> (ICN)	Informal clade name for the group containing <i>Cyclanorbis senegalensis</i> (Dumeril and Bibron, 1835) and <i>Cyclanorbis elegans</i> (Gray, 1869). Not formally defined here because of ambiguity regarding the relationship of <i>C. senegalensis</i> to other members of <i>Cyclanorbininae</i> .
<i>Cycloderma</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Cycloderma frenatum</i> (Peters, 1854) and <i>Cycloderma aubryi</i> (Dumeril, 1856).
<i>Lissemys</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Lissemys punctata</i> (Lacépède, 1788) and <i>Lissemys scutata</i> (Peters, 1868).
<i>Trionychinae</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Trionyx triunguis</i> (Forskål, 1775), <i>Apalone ferox</i> (Schneider, 1783), and <i>Amyda cartilaginea</i> (Boddaert, 1770).
<i>Gigantaestuarochelys</i> (NCN)	Refers to the crown clade arising from the last common ancestor of <i>Trionyx triunguis</i> (Forskål, 1775), <i>Chitra indica</i> (Gray, 1844), and <i>Pelochelys bibroni</i> (Owen, 1853).
<i>Chitraina</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Chitra indica</i> (Gray, 1844) and <i>Pelochelys bibroni</i> (Owen, 1853).
<i>Chitra</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Chitra indica</i> (Gray, 1844), <i>Chitra chitra</i> (Nutaphand, 1986), and <i>Chitra vandijki</i> (McCord and Pritchard, 2002).
<i>Pelochelys</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Pelochelys bibroni</i> (Owen, 1853) and <i>Pelochelys cantorii</i> (Gray, 1864).
<i>Apalonina</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Rafetus euphraticus</i> (Daudin, 1802) and <i>Apalone ferox</i> (Schneider, 1783).
<i>Rafetus</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Rafetus euphraticus</i> (Daudin, 1802) and <i>Rafetus swinhoei</i> (Gray, 1873).
<i>Apalone</i> (CCN)	Refers to the crown clade arising from the last common ancestor of <i>Apalone ferox</i> (Schneider, 1783) and <i>Apalone mutica</i> (Le Sueur, 1827).
<i>Amydona</i> (NCN)	Refers to the crown clade arising from the last common ancestor of <i>Amyda cartilaginea</i> (Boddaert, 1770), <i>Pelodiscus sinensis</i> (Wiegman, 1835), and <i>Aspideretini formosa</i> (Gray, 1869).
<i>Aspideretini</i> (CCN)	Refers to the crown clade arising from the last common ancestor of " <i>Nilssonina</i> " <i>formosa</i> (Gray, 1869), " <i>Aspideretes</i> " <i>gangeticus</i> (Cuvier, 1831), " <i>Aspideretes</i> " <i>hurum</i> (Gray, 1831), " <i>Aspideretes</i> " <i>leithii</i> (Gray, 1872), " <i>Aspideretes</i> " <i>nigricans</i> (Anderson, 1875).

of *Apalone ferox* and *Rafetus euphraticus*. Meylan (1987) erected the name *Apalonina* for this biogeographically disparate clade, which was first recognized by Loveridge and Williams (1957) based on the reduced or absent 8th pleurals (Meylan 1987, character 48).

The current generic taxonomy of softshell turtles provides several interesting puzzles. Of the eight currently recognized genera that are not monotypic, six appear to be unquestionably monophyletic groups. These include *Apalone* (3 species), *Chitra* (3 species), *Cycloderma* (2 species), *Rafetus* (2 species), *Lissemys* (2 species), and *Pelochelys* (2 species). The other two genera, *Cyclanorbis* and *Aspideretes*, are paraphyletic in many of our analyses; *Cyclanorbis* with respect to *Cycloderma* and *Lissemys*, and *Aspideretes* with respect to *Nilssonina formosa*. We do not recommend changing the content of *Cyclanorbis* until further studies clarify its content and relationships to other flap-shelled turtles. In the case of *Aspideretes*, the older genus name *Nilssonina* (Gray, 1872) has precedence over *Aspideretes* (Hay, 1903 [but commonly reported as 1904]), thus with traditional rank-based taxonomy with the mandatory rank of genus, changing the taxonomy to re-

flect our best current estimate of phylogeny of these turtles requires both changing the genus name *Aspideretes* to *Nilssonina* and changing the spelling of the species names of the four species of *Aspideretes* to the feminine gender to match *Nilssonina*. We are reluctant to suggest such sweeping changes given that there is some ambiguity in our analyses regarding these relationships. Our preferred solution, given our current level of knowledge, is to abandon the current generic names, and refer all five species to a higher clade, *Aspideretini*, defined as the clade arising from the last common ancestor of these five species (Table 3). This unranked clade definition would provide recognition the close relationship of *formosa* and the four species currently classified as *Aspideretes*, and would remain taxonomically stable under changing ideas of the relationships within the group. Although this is our recommendation, we recognize that some biologists will be unwilling to accept a taxonomy, which does not include a genus name. For them, we recommend that the current binomials be maintained, but with the tacit recognition that the genus names are placeholders not intended to convey any phylogenetic information until there is

convincing evidence on the relationships among the five species of the crown clade *Aspideretini*.

The remaining five currently recognized genera are monotypic. These monotypic genera were resurrected to maintain a consistent cladistic taxonomy and to recognize morphological uniqueness of species (Meylan, 1987). Our molecular results support the idea that these monotypic genera represent old lineages with a long history of independent evolution (the minimum mtDNA sequence divergence in our data set among taxa currently classified in different genera is 9.03% between "*Nilsso-nia*" *formosa* and "*Aspideretes*" *hurum*). However, the idea of anything approaching a consistent "generic level of differentiation" for either molecular or morphological characters has never been seriously considered, even within the turtle systematic community. Thus, it is not clear what, if any, useful information monotypic genus names convey (Spinks et al., 2004). Monotypic genera (or any other higher category name) also fail to convey unique phylogenetic information, making them literally redundant under phylogenetic taxonomy (de Queiroz and Gauthier, 1994; de Queiroz and Cantino, 2001). Although, in general we do not favor the monotypic genus, in this case we feel that there is little to gain from collapsing the current widely used names. Four of these monotypic genera (*Amyda*, *Dogania*, *Palea*, and *Pelodiscus*) are members of the highly pectinate *Amydona*. There is no obvious ecological, morphological, or phylogenetic distinction within this group that we feel merits special taxonomic recognition, and lumping the entire group in a single genus name would convey no more (or less) phylogenetic information that is present in the unranked clade *Amydona*. The fifth monotypic genus, *Trionyx*, is a member of the *Gigantaestuarochelys* and sole sister group of the *Chitra* (*Chitra*+*Pelochelys*). *Chitra* (Gray, 1864) and *Pelochelys* (Gray, 1864) are the only two names that have remained stable and consistent over the past century, but the older *Trionyx* (Geoffroy, 1809a, 1809b) would have precedence over these two well-known names. Subsuming *Chitra* and *Pelochelys* into *Trionyx* would convey no more phylogenetic information than is present in the unranked clade name *Gigantaestuarochelys* and, ironically, would be completely at odds with former usage of the name *Trionyx* to refer to all *Trionychines* except for *Chitra* and *Pelochelys*.

Further motivation for the maintenance of current generic names is the recent discovery of considerable unrecognized diversity within taxa, which were formerly considered to be monotypic (Engstrom et al., 2002; Engstrom and McCord, 2002; McCord and Pritchard, 2002). The high sequence divergence in our molecular data confirms that several recently recognized species including *Lissemys punctata* and *L. scutata* (Webb, 1982: 10.5% mtDNA), *Pelochelys bibroni* and *P. cantorii* (Webb, 1995: 6.5%), *Chitra indica*, *C. chitra* (Nutaphand, 1986: 6.6%), and *C. vandijki* (McCord and Pritchard, 2002: 4.4%) represent divergent evolutionary lineages that are consistent with species status (Wiens and Penkrot, 2002). This suggests that species diversity of softshell turtles may be much greater than presently recognized

(Engstrom et al., 2002). If this is true, then the monotypic genus names, which at present seem superfluous, may at some point in the near future be able to convey useful phylogenetic information about as yet undescribed, and unnamed, phylogenetic diversity.

#### *Analytical Strategies for Extracting Phylogenetic Signal from Saturated mtDNA Data*

The theoretical, empirical, and computational challenges of recovering accurate phylogenetic information from DNA sequence data have proven very substantial (Sanderson and Shaffer, 2002). This seems particularly true for studies using mtDNA to resolve deep phylogenetic relationships. These studies often result in poorly resolved, starburst-like phylogenies consistent with an ancient rapid diversification of lineages (Shaffer et al., 1997; Mahoney, 2001). In other cases mtDNA consistently support anomalous relationships that are widely believed to be incorrect based on strong morphological and/or fossil data (Naylor and Brown, 1998; Garcia-Machado et al., 1999; Wiens and Hollingsworth, 2000). In this study we have seen instances of each type of behavior. Equally weighted parsimony analyses of mitochondrial genes resulted in a poorly resolved phylogeny consistent with rapid early divergence. Mitochondrial data also consistently places *Rafetus euphraticus* as a member of the Asian clade whereas nuclear and morphological data clearly place *Rafetus* as sister to the North American *Apalone*.

A possible solution to this problem with mitochondrial data could be simply to seek better data (Sanderson and Shaffer, 2002) in the form of nuclear sequence data (e.g., this study; Prychitko and Moore, 1997; Georges et al., 1999; Baker et al., 2001; Matthee et al., 2001; Springer et al., 2001a), morphological data (Shaffer et al., 1997; Wiens and Reeder, 1995, 1997; Wiens, 1998), or fossil data (Shaffer et al., 1997; Springer et al., 2001b; Soltis et al., 2002). Although "simply" collecting better data may be the most desirable solution, this is not always possible (due to a lack of markers) or economically feasible. We have also seen from our data that including more nuclear and morphological data that are less homoplasious does not automatically lead to stronger conclusions (Appendix 1). This, combined with the unfortunate possibility that many studies that were initially done with mtDNA will not be replicated with more or better genes, has led many systematists to attempt to rescue mtDNA on two fronts: (1) more complete taxon sampling (whether collecting the data themselves or harvesting data from databases), and (2) more sophisticated analyses. Increased taxon sampling can reduce error and increase accuracy of phylogenetic inference (Hillis, 1996, 1998; Källersjö et al., 1999; Pollock et al., 2002; Zwickl and Hillis, 2002), and is a reasonable approach when possible. However, increased sampling may be impossible because many groups, including softshell turtles, simply contain relatively few living taxa. In these cases the only option available is a more critical analyses of the available data. We examined two main categories of

critical analyses of these data: (1) weighted parsimony, in which some attempt is made to identify, and then down-weight or eliminate, potentially misleading data; and (2) model-based approaches including maximum likelihood and Bayesian phylogenetic inference, which compensate for homoplasy by identifying and utilizing realistic models of character evolution.

*Parsimony: weighting and weeding.*—Weighted parsimony has led to mixed positive (Chippindale and Wiens, 1994; Allard and Carpenter, 1996; Voelker and Edwards, 1998; Barker and Lanyon, 2000) and negative (Källersjö et al., 1999; Broughton et al., 2000; Frost et al., 2001) results. The two mitochondrial protein-coding genes used in this study show many of the classical signs of excessive homoplasy attributable to saturation with multiple substitutions, and fail to recover relationships that are very strongly supported by non-homoplasious characters in the nuclear intron. Our use of saturation plots to identify highly homoplasious, potentially misleading data may be problematic. Kjer et al. (2001) highlight the flaws of the concept of “saturation” by pointing out that as generally discussed, “saturation” is a distance-based concept, but in character-based phylogeny methods (e.g., parsimony or likelihood), “saturated” data may still provide meaningful phylogenetic information, especially with adequate taxonomic sampling (Hillis, 1996, 1998; Yang and Goldman, 1997; Graybeal, 1998; Pollock et al., 2002; Zwickl and Hillis, 2002). This is because homoplasy in different parts of the tree can be efficiently isolated (Swofford et al., 1996; Kjer et al., 2001). In our case (and possibly many others) in which extensive taxon sampling is simply not an option, we found that saturation plots were a useful heuristic device to identify, candidate data for weighted parsimony.

Our results echo those of Broughton et al. (2000) in that parsimony analyses of mitochondrial protein-coding genes singly and combined failed to recover all of the well-supported deep nodes in the softshell turtles phylogeny when equally weighted or with any differential weighting scheme. However, all of the parsimony weeding and weighting techniques used here increased bootstrap support for most of the difficult nodes compared to equally weighted parsimony. The methods employing ti/tv weighting using  $1/\kappa$ , and weeding of saturated sites, both of which retain information from transitions, were most effective, and almost always resulted in increased overall tree support and support for key deep nodes. Although neither of these two methods was obviously superior to the other, our preferred scheme is  $1/\kappa$  transition weighting because it retains potentially valuable information from all transitions, without artificially increasing the weight of transversions in the mtDNA data. This is particularly valuable when combining noisy mtDNA data with nuclear and morphological data sets containing fewer informative characters.

*Model-based approaches: likelihood and Bayesian analyses.*—Maximum-likelihood analyses of both weeded and unweeded data performed better than any parsimony technique that we evaluated, according to our heuristic measure of summed overall bootstrap

support. This was true for both measures of BP support across the entire tree and for the six difficult nodes. In contrast to parsimony, weeded maximum-likelihood analysis did not show a marked improvement over analysis of the unweeded data. These homoplasious data contain useful information, which can be accessed by using appropriate models of DNA evolution. Maximum-likelihood analysis is more effective when considering this information than when these data are completely eliminated. Although maximum-likelihood analyses of mtDNA data recovered many deep nodes with higher levels of bootstrap support than parsimony, no maximum-likelihood analysis of mtDNA alone was able to recover the sister relationship of *Rafetus* and *Apalone*. Thus, exclusive reliance on mtDNA would have left us with a very different view the evolutionary history of this biogeographically fascinating group. As with previous studies, using partitioned maximum-likelihood (DeBry, 1999; Wilgenbusch and De Queiroz, 2000; Caterino et al., 2001), we found a substantial improvement in likelihood scores for partitioned likelihood compared to global models. Partitioned maximum-likelihood also served well as a means to highlight areas of conflict in our data set. In the case of *Rafetus*, partitioned maximum-likelihood showed that even with a parameter-rich model of molecular evolution, maximum likelihood can be susceptible to long-branch attraction (Felsenstein, 1978), by highlighting the support from 3rd positions for the nonmonophyly of *Rafetus* and *Apalone*. This as an indication that the conflict among fast and slowly evolving characters in these ancient lineages is caused by homoplasy in the fast evolving characters.

Bayesian analyses is a powerful tool for inference (Huelsenbeck et al., 2001), and may be the only practical means by which to apply model based phylogenetic inference to large data sets (Leaché and Reeder, 2002) and data sets containing both molecular and morphological data. Our Bayesian analyses of molecular data recovered the six deep nodes with high posterior probabilities without sacrificing resolution of shallow nodes. Bayesian analyses of combined molecular and morphological data were largely congruent with analyses of molecular data alone and with parsimony-based analyses of the combined morphological and molecular data. However, the Bayesian support for the contested relationships within *Apalone* was much higher than in molecular data alone or parsimony analyses (100 PP versus 76). This is not surprising in light of simulation studies showing the high sensitivity of Bayesian analyses to small amounts of signal in the data (Alfaro et al., 2003). As a cautionary note, however, Buckley et al. (2002) and Erixon et al. (2003) have shown that model mis-specification can also result in overconfidence in the results of Bayesian analyses. Because maximum-likelihood methods for modeling morphological character evolution (Lewis, 2001) are relatively new and have not been extensively tested empirically, it still seems advisable to continue to employ both Bayesian and parsimony methods such as those outlined here for mixed molecular and morphological data sets.

## CONCLUSION

We set out to provide a working hypothesis for the phylogeny of softshell turtles and to examine strategies for analyzing saturated mtDNA data in deep phylogenetic problems. We found that mtDNA provided good resolution for shallow divergence, but experienced long branch-related problems with deep divergences. Weeded and weighted parsimony and model-based techniques were only partially able to mitigate these problems. Only by the addition of largely nonhomoplasious data from a nuclear intron were we able to confidently recover many deep nodes in the softshell turtle phylogeny. The R35 intron used here provided remarkably good resolution for deep nodes, but contained limited information regarding recent divergences. By combining the intron with the more variable mtDNA data we were able to confidently reconstruct these shallow nodes as well. Using just one of these markers would not have provided a well-resolved phylogeny, even with the most sophisticated analytical strategies we could find in the literature. Analysis of these heterogeneous data using partitioned models of sequence evolution removes the problems associated with combining data and appears to be the most logical way to analyze heterogeneous data. By further combining this DNA data set with a complementary morphological data set, we have been able to provide a well-supported, completely sampled phylogenetic hypothesis for softshell turtles, and have identified a small number of ambiguous relationships that will require further clarification. We hope that this phylogeny can be used to frame future research on this group and that the classification built on this phylogeny will facilitate effective communication of that research.

## ACKNOWLEDGMENTS

We thank the many people who have contributed live specimens to the McCord living turtle collection over the years and to Patrick Baker for his curation of the collection. Additional samples were provided by Paul Moler (Florida Game and Fresh Water Fish Commission) and Ertan Taskavak (Ege University). Funding was provided to TNE by an NSF predoctoral fellowship and by grants from the Center for Population Biology (UCD), Jastro-Shields Fund (UCD), and UCD Humanities Research Program. The Wildlife Conservation Society supported fieldwork in Turkey. This material was based on funding from NSF grants (DEB-9727161 and DEB-0213155) to HBS. Peter Meylan provided the initial motivation for this study, and contributed invaluable assistance and insight throughout the project, and Mathew Fujita developed the R35 intron primers. Kevin de Queiroz and Walter Joyce provided advice regarding rank-free classification. Mike Sanderson, John Gillespie, and members of the Shaffer lab at UCD and the Wiens lab at SBU provided comments on earlier versions of this manuscript. This work would not have been possible without the continued support of Kristine Mazzei.

## REFERENCES

- Alfaro, M. E., S. Zoller, and F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Mol. Biol. Evol.* 20:255–266.
- Allard, M. W., and J. M. Carpenter. 1996. On weighting and congruence. *Cladistics* 12:183–198.
- Arctander, P. 1995. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proc. R. Soc. Lond. B. Biol. Sci.* 262:13–19.
- Arevalo, E., S. K. Davis, and J. W. Sites. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. *Syst. Biol.* 43:387–418.
- Armstrong, M. H., E. L. Braun, and R. T. Kimball. 2001. Phylogenetic utility of avian ovomucoid intron G: A comparison of nuclear and mitochondrial phylogenies in galliformes. *Auk* 118:799–804.
- Baker, R. H., G. S. Wilkinson, and R. DeSalle. 2001. Phylogenetic utility of different types of molecular data used to infer evolutionary relationships among stalk-eyed flies (Diopsidae). *Syst. Biol.* 50:87–105.
- Barker, F. K., and S. M. Lanyon. 2000. The impact of parsimony weighting schemes on inferred relationships among toucans and Neotropical barbets (Aves: Piciformes). *Mol. Phylogenet. Evol.* 15:215–234.
- Bensasson, D., D. X. Zhang, and G. M. Hewitt. 2000. Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Mol. Biol. Evol.* 17:406–415.
- Birks, S. M., and S. V. Edwards. 2002. A phylogeny of the megapodes (Aves: Megapodiidae) based on nuclear and mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 23:408–421.
- Boulenger, G. A. 1889. Catalogue of the chelonians, rhynchocephalians, and crocodiles in the British Museum (Natural History), new edition. British Museum of Natural History, London.
- Broughton, R. E., S. E. Stanley, and R. T. Durrett. 2000. Quantification of homoplasy for nucleotide transitions and transversions and a reexamination of assumptions in weighted phylogenetic analyses. *Syst. Biol.* 49:617–627.
- Brown, W. M., M. George Jr, and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U. S. A.* 76:1967–1971.
- Buckley, T. R., and C. W. Cunningham. 2002. The effects of nucleotide substitution model assumptions on estimates of nonparametric bootstrap support. *Mol. Biol. Evol.* 19:394–405.
- Caterino, M. S., R. D. Reed, M. M. Kuo, and F. A. H. Sperling. 2001. A partitioned likelihood analyses of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). *Syst. Biol.* 50:106–127.
- Chippindale, P. T., and J. J. Wiens. 1994. Weighting, partitioning, and combining characters in phylogenetic analyses. *Syst. Biol.* 43:278–287.
- Chkhikvadze, V. M. 2000. Systematic status of extinct soft-shelled turtles of the subgenus *Rafetus* (*Axestemys*). *J. Vert. Paleo.* 20:35A.
- Conant, G. C., and P. O. Lewis. 2001. Effects of nucleotide composition bias on the success of the parsimony criterion in phylogenetic inference. *Mol. Biol. Evol.* 18:1024–1033.
- de Broin, F. 1977. Contribution à l'étude des chéloniens, chéloniens continentaux de Crétacé et du Tertiaire de France. *Mus. Natl. Hist. Nat.*, Mémoires ser. C 38:1–366.
- de Queiroz, K., and P. D. Cantino. 2001. Phylogenetic nomenclature and the PhyloCode. *Bull. Zool. Nom.* 58:254–271.
- de Queiroz, K., and J. Gauthier. 1994. Toward a phylogenetic system of biological nomenclature. *Trends Ecol. Evol.* 9:27–31.
- DeBry, R. W. 1999. Maximum likelihood analyses of gene-based and structure-based process partitions, using mammalian mitochondrial genomes. *Syst. Biol.* 48:286–299.
- Degli, E. M., V. S. De, M. Crimi, A. Ghelli, T. Patarnello, and A. Meyer. 1993. Mitochondrial cytochrome *b*: Evolution and structure of the protein. *Biochim. Biophys. Acta* 1143:243–271.
- DeWoody, J. A., R. K. Chesser, and R. J. Baker. 1999. A translocated mitochondrial cytochrome *b* pseudogene in voles (Rodentia: Microtus). *J. Mol. Evol.* 48:380–382.
- Engstrom, T. N., and W. P. McCord. 2002. Molecular support for the taxonomic conclusions of McCord and Pritchard (2002), regarding *Chitra*. *Hamadryad* 27:57–61.
- Engstrom, T. N., H. B. Shaffer, and W. P. McCord. 2002. Phylogenetic diversity of endangered and critically endangered southeast Asian softshell turtles (Trionychidae: *Chitra*). *Biol. Conserv.* 104:173–179.
- Erixon, P., B. Sennblad, T. Britton, and B. Oxelman. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Syst. Biol.* 52:665–673.
- Ernst, C. H., and R. W. Barbour. 1989. *Turtles of the world*. Smithsonian Institution Press, Washington, DC.
- Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401–410.



- Friedel, R. H., J. Stubbush, Y. A. Barde, and H. Schnuerch. 2001. A novel 7-transmembrane receptor expressed in nerve growth factor-dependent sensory neurons. *Mol. Cell. Neurosci.* 17:31–40.
- Friesen, V. L., B. C. Congdon, M. G. Kidd, and T. P. Birt. 1999. Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. *Mol. Ecol.* 8:2147–2149.
- Frost, D. R., M. T. Rodrigues, T. Grant, and T. A. Titus. 2001. Phylogenetics of the lizard genus *Tropidurus* (Squamata: Tropiduridae: Tropidurinae): Direct optimization, descriptive efficiency, and sensitivity analyses of congruence between molecular data and morphology. *Mol. Phylogenet. Evol.* 21:352–371.
- Fujita, M. K., T. N. Engstrom, D. E. Starkey, and H. B. Shaffer. 2004. Turtle phylogeny: Insights from a novel nuclear intron. *Mol. Phylogenet. Evol.* 31:1031–1040.
- Gaffney, E. S. 1979a. Fossil trionychids of Australia. *J. Paleont.* 54:1354–1360.
- Gaffney, E. S. 1979b. Description of a large trionychid turtle shell from the Eocene Bridger Formation of Wyoming. *Contrib. Geol.* 17:53–57.
- Gaffney, E. S., and P. A. Meylan. 1988. A phylogeny of turtles. Pages 157–219 in *The phylogeny and classification of tetrapods* (M. J. Benton, ed.). Clarendon Press, Oxford, England.
- Garcia-Machado, E., M. Pempera, N. Dennebouy, M. Oliva-Suarez, J. C. Mounolou, and M. Monnerot. 1999. Mitochondrial genes collectively suggest the paraphyly of crustacea with respect to insecta. *J. Mol. Evol.* 49:142–149.
- Geoffroy Saint-Hilaire, M. 1809a. Sur les tortues molles, nouveau genre sous le nom de *Trionyx* et sur la formation des carapaces. *Ann. Mus. Hist. Natur. Paris* 14:1–20.
- Geoffroy Saint-Hilaire, M. 1809b. Mémoire sur les tortues molles. *Bulletin des Sciences, par la Société Philomatique de Paris*, 1809:363–367.
- Georges, A., J. Birrell, K. M. Saint, W. McCord, and S. C. Donnellan. 1999. A phylogeny for side-necked turtles (Chelonio: Pleurodira) based on mitochondrial and nuclear gene sequence variation. *Biol. J. Linn. Soc.* 67:213–246.
- Gray, J. E. 1864. Revision of the species of Trionychidae, found in Asia and Africa, with description of some new species. *Proc. Zool. Soc. Lond.* 1864:76–98.
- Gray, J. E. 1872. On Indian mud-tortoises (*Trionyx*). *Ann. Mag. Nat. Hist.* 4:326–340.
- Gray, J. E. 1873. Notes on mud-tortoises (*Trionyx*: Geoffroy) and on the skulls of the different kinds. *Proc. Zool. Soc. Lond.* 1873:38–72.
- Graybeal, A. 1993. The phylogenetic utility of cytochrome *b*: Lessons from bufonid frogs. *Mol. Phylogenet. Evol.* 2:256–269.
- Graybeal, A. 1994. Evaluating the phylogenetic utility of genes: A search for genes informative about deep divergences among vertebrates. *Syst. Biol.* 43:174–193.
- Graybeal, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* 47:9–17.
- Greenhalgh, P., C. E. M. Olesen, and L. A. Steiner. 1993. Characterization and expression of recombination activating genes (RAG-1 and RAG-2) in *Xenopus laevis*. *J. Immunol.* 151:3100–3110.
- Griffiths, C. S. 1997. Correlation of functional domains and rates of nucleotide substitution in cytochrome *b*. *Mol. Phylogenet. Evol.* 7:352–365.
- Groth, J. G., and G. F. Barrowclough. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol. Phylogenet. Evol.* 12:115–123.
- Gu, X., and W.-H. Li. 1998. Estimation of evolutionary distances under stationary and nonstationary models of nucleotide substitution. *Proc. Natl. Acad. Sci. U. S. A.* 95:5899–5905.
- Hay, O. P. 1903. On the existing genera of the Trionychidae. *Proc. Am. Phil. Soc.* 42:268–274.
- Hedges, S. B., and L. L. Poling. 1999. A molecular phylogeny of reptiles. *Science* 283:998–1001.
- Hillis, D. M. 1996. Inferring complex phylogenies. *Nature* 383:130–131.
- Hillis, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Syst. Biol.* 47:3–8.
- Hillis, D. M., B. K. Mable, and C. Moritz. 1996. Applications of molecular systematics: The state of the field and a look to the future. Pages 515–543 in *Molecular systematics*, second edition. (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer Associates, Sunderland, Massachusetts.
- Huelsenbeck, J. P., D. M. Hillis, and R. Nielsen. 1996. A likelihood-ratio test of monophyly. *Syst. Biol.* 45:546–558.
- Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
- Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294:2310–2314.
- Hummel, K. 1929. Die fossilen Weichschildkroten (Trionychia). Eine morphologisch-systematische und stammengeschichtliche Studie. *Geol. Palaeont. Abh.* 16:359–487.
- International Code of Zoological Nomenclature. 1999. Fourth Edition. International Trust for Zoological Nomenclature, London.
- Iverson, J. B. 1992. A revised checklist with distribution maps of the turtles of the world. Privately published, Richmond, Indiana.
- Irwin, D. M., T. D. Kocher, and A. C. Wilson. 1991. Evolution of the cytochrome *b* gene of mammals. *J. Mol. Evol.* 32:128–144.
- Joyce, W. G., J. F. Parham, and J. A. Gauthier. 2004. Developing a protocol for the conversion of rank-based taxon names to phylogenetically defined clade names, as exemplified by turtles. *J. Paleontol.* 78:989–1013.
- Källersjö, M., V. A. Albert, and J. S. Farris. 1999. Homoplasy increases phylogenetic structure. *Cladistics* 15:91–93.
- Källersjö, M., J. S. Farris, M. W. Chase, B. Bremer, M. F. Fay, C. J. Humphries, G. Petersen, O. Seberg, and K. Bremer. 1998. Simultaneous parsimony jackknife analyses of 2538 rbcL DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants. *Pl. Syst. Evol.* 213:259–287.
- Kordikova, E. G. 1991. Review of fossil trionychids of Kazakhstan (in Russian). *And. Geol. Mineral. Nauk. Tbilisi* 1991:1–13.
- Kjer, K. M., R. J. Blahnik, and R. W. Holzenthal. 2001. Phylogeny of Trichoptera (Caddisflies): Characterization of signal and noise within multiple datasets. *Syst. Biol.* 50:781–816.
- Kumazawa, Y., and M. Nishida. 1999. Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: Statistical evidence for Archosaurian affinity of turtles. *Mol. Biol. Evol.* 16:784–792.
- Leaché, A. D., and T. W. Reeder. 2002. Molecular systematics of the Eastern Fence Lizard (*Sceloporus undulatus*): A comparison of parsimony, likelihood, and Bayesian approaches. *Syst. Biol.* 51:44–68.
- Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50:913–925.
- Loveridge, A., and E. E. Williams. 1957. Revision of the African tortoises and turtles of the suborder Cyrtodira. *Bull. Mus. Comp. Zool.* 115:163–557.
- Lü, X. M., Y. X. Fu, and Y. P. Zhang. 2002. Evolution of mitochondrial cytochrome *b* pseudogene in genus *Nycticebus*. *Mol. Biol. Evol.* 19:2337–2341.
- Lydekker, R. 1889. Catalogue of the fossil Reptilia and Amphibia in the British Museum (Natural History), Part III (containing the order Chelonio). Taylor and Francis, London.
- McCord, W. P., and P. C. H. Pritchard. 2002. A review of the softshell turtles of the genus *Chitra*, with the description of new taxa from Myanmar and Indonesia (Java). *Hamadryad* 27:11–56.
- Macey, J. R., A. Larson, N. B. Ananjeva, Z. Fang, and T. J. Papenfuss. 1997b. Two novel gene orders and the role of light-strand replication in the rearrangement of the vertebrate mitochondrial genome. *Mol. Biol. Evol.* 14:91–104.
- Macey, J. R., A. Larson, N. B. Ananjeva, and T. J. Papenfuss. 1997a. Evolutionary shifts in three major structural features of the mitochondrial genome among iguanian lizards. *J. Mol. Evol.* 44:660–674.
- Maddison, W. P., and D. R. Maddison. 2003. Mesquite: A modular system for evolutionary analyses. Version 0.996. <http://mesquiteproject.org>.
- Mahoney, M. J. 2001. Molecular systematics of *Plethodon* and *Aneides* (Caudata: Plethodontidae: Plethodontini): Phylogenetic analyses of an old and rapid radiation. *Mol. Phylogenet. Evol.* 18:174–188.
- Matthee, C. A., J. D. Burzlaff, J. F. Taylor, and S. K. Davis. 2001. Mining the mammalian genome for artiodactyl systematics. *Syst. Biol.* 50:367–390.
- Meyer, A. 1994. Shortcomings of the cytochrome *b* gene as a molecular marker. *Trends Ecol. Evol.* 9:278–280.
- Meylan, P. A. 1987. The phylogenetic relationships of soft-shelled turtles (family Trionychidae). *Bulletin of the American Museum of Natural History* 186:1–101.

- Mindell, D. P., M. D. Sorenson, and D. E. Dimcheff. 1998. An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.* 15:1568–1571.
- Mindell, D. P., M. D. Sorenson, D. E. Dimcheff, M. Hasegawa, J. C. Ast, and T. Yuri. 1999. Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. *Syst. Biol.* 48:138–152.
- Naylor, G. J. P., and W. M. Brown. 1998. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. *Syst. Biol.* 47:61–76.
- Nie, L. W., X. X. Shan, M. Wang, C. W. Guo, and X. X. Lu. 2001. The conservative region sequence analyses of four Sox Genes in the *Trionyx sinensis*. *Acta Hydrobiol. Sinica* 25:245–250.
- Nutaphand, W. 1986. "Manlai," the world's biggest softshell turtle. (In Thai.) *Thai Zool. Mag.* 1:64–70.
- Palumbi, S. R. 1996. Nucleic acids II: The polymerase chain reaction. Pages 205–247 in *Molecular systematics*, second edition. (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer Associates, Sunderland, Massachusetts.
- Plummer, M. V. 2001. Testudines: *Trionyx muticus* (Smooth softshell turtle). *Foraging. Herpetol. Rev.* 32:249–250.
- Pollock, D. D., D. J. Zwickl, J. A. McGuire, and D. M. Hillis. 2002. Increased taxon sampling is advantageous for phylogenetic inference. *Syst. Biol.* 51:664–671.
- Posada, D. 2001. Modeltest, version 3.06. Department of Zoology, Brigham Young University, Provo, Utah.
- Pritchard, P. C. H. 2001. Observations on body size, sympatry, and niche divergence in softshell turtles (Trionychidae). *Chelonian Conserv. Biol.* 4:5–27.
- Prychitko, T. M., and W. S. Moore. 1997. The utility of DNA sequences of an intron from the beta-fibrinogen gene in phylogenetic analyses of woodpeckers (Aves: Picidae). *Mol. Phylogenet. Evol.* 8:193–204.
- Prychitko, T. M., and W. S. Moore. 2000. Comparative evolution of the mitochondrial cytochrome *b* gene and nuclear beta-fibrinogen intron 7 in woodpeckers. *Mol. Biol. Evol.* 17:1101–1111.
- Prychitko, T. M., and W. S. Moore. 2003. Alignment and phylogenetic analyses of *b*-fibrinogen intron 7 sequences among avian orders reveal conserved regions within the intron. *Mol. Biol. Evol.* 20:762–771.
- Romer, A. S. 1968. Osteology of the reptiles. University of Chicago Press, Chicago.
- Saint, K. M., C. C. Austin, S. C. Donnellan, and M. N. Hutchinson. 1998. C-mos, a nuclear marker useful for squamate phylogenetic analyses. *Mol. Phylogenet. Evol.* 10:259–263.
- Sanderson, M. J., and M. J. Donoghue. 1996. The relationship between homoplasy and confidence in a phylogenetic tree. Pages 67–89 in *Homoplasy: The recurrence of similarity in evolution* (M. J. Sanderson and L. Hufford, eds.). Academic Press, San Diego.
- Sanderson, M. J., and L. Hufford (eds.). 1996. Homoplasy: The recurrence of similarity in evolution. Academic Press, San Diego.
- Sanderson, M. J., and H. B. Shaffer. 2002. Troubleshotting molecular phylogenetic analyses. *Ann. Rev. Ecol. Syst.* 33:49–72.
- Schulte, J. A. III, J. P. Valladares, and A. Larson. 2003. Phylogenetic relationships within Iguanidae inferred using molecular and morphological data and a phylogenetic taxonomy of iguanian lizards. *Herpetologica* 59:399–419.
- Shaffer, H. B., J. M. Clark, and F. Kraus. 1991. When molecules and morphology clash: A phylogenetic analyses of the North American ambystomatid salamanders (Caudata: Ambystomatidae). *Syst. Zool.* 40:284–303.
- Shaffer, H. B., P. A. Meylan, and M. L. McKnight. 1997. Tests of turtle phylogeny: Molecular, morphological, and paleontological approaches. *Syst. Biol.* 46:235–268.
- Soltis, P. S., D. E. Soltis, V. Savolainen, P. R. Crane, and T. G. Barraclough. 2002. Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil data and evidence for molecular living fossils. *Proc. Natl. Acad. Sci. U. S. A.* 99:4430–4435.
- Spinks, P. Q., G. B. Pauly, J. J. Crayon, and H. B. Shaffer. 2003. Survival of the Western Pond Turtle (*Emys marmorata*) in an Urban California Environment. *Biol. Conserv.* 113:257–267.
- Spinks, P. Q., H. B. Shaffer, J. B. Iverson, and W. P. McCord. 2004. Phylogenetic hypothesis for the turtle family Geoemydidae. *Mol. Phylogenet. Evol.* 32:164–182.
- Springer, M. S., R. W. DeBry, C. Douady, H. M. Amrine, O. Madsen, W. W. de Jong, and M. J. Stanhope. 2001a. Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. *Mol. Biol. Evol.* 18:132–143.
- Springer, M. S., E. C. Teeling, O. Madsen, M. J. Stanhope, and W. W. de Jong. 2001b. Integrated fossil and molecular data reconstruct bat echolocation. *Proc. Natl. Acad. Sci. U. S. A.* 98:6241–6246.
- Starkey, D. E. 1997. Molecular systematics and biogeography of the new world turtle genera *Trachemys* and *Kinosternon*. Texas A&M University, College Station, Texas.
- Swofford, D. L. 2002. PAUP\* Phylogenetic Analysis Using Parsimony (\*and Other Methods), version 4. Sinauer Associates, Sunderland, Massachusetts.
- Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. Pages 407–514 in *Molecular systematics*, second edition. (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer Associates, Sunderland, Massachusetts.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL-X windows interface: Flexible strategies for multiple sequence alignment aided by quality analyses tools. *Nucleic Acids Res.* 25:4876–4882.
- van Dijk, P. P., B. L. Stuart, and A. G. J. Rhodin (eds.). 2000. Asian Turtle Trade: Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia. Chelonian Research Foundation, Lundberg, Massachusetts.
- Voelker, G., and S. V. Edwards. 1998. Can weighting improve bushy trees? Models of cytochrome *b* evolution and the molecular systematics of pipits and wagtails (Aves: Motacillidae). *Syst. Biol.* 47:589–603.
- Webb, R. G. 1982. Taxonomic notes concerning the trionychid turtle *Lissemys punctata* (Lacépède). *Amphibia-Reptilia* 3:179–184.
- Webb, R. G. 1995. Redescription and neotype designation of *Pelochelys bibroni* from southern New Guinea (Testudines: Trionychidae). *Chelonian Conserv. Biol.* 1:301–310.
- Weisrock, D. W., and F. J. Janzen. 2000. Comparative molecular phylogeography of North American softshell turtles (Apalone): Implications for regional and wide-scale historical evolutionary forces. *Mol. Phylo. Evol.* 14:152–164.
- White, P. S., and L. D. Densmore, II. 1992. Mitochondrial DNA isolation. Pages 29–58 in *Practical approach series: Molecular genetic analysis of populations*. (A. R. Hoelzel, ed.). Oxford University Press, Oxford, England.
- Wiens, J. J. 1998. Does adding characters with missing data increase or decrease phylogenetic accuracy? *Syst. Biol.* 47:625–640.
- Wiens, J. J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. *Syst. Biol.* 52:528–538.
- Wiens, J. J., and B. D. Hollingsworth. 2000. War of the Iguanas: Conflicting molecular and morphological phylogenies and long-branch attraction in iguanid lizards. *Syst. Biol.* 49:143–159.
- Wiens, J. J., and T. A. Penkrot. 2002. Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Syst. Biol.* 51:69–91.
- Wiens, J. J., and T. W. Reeder. 1995. Combining data sets with different numbers of taxa for phylogenetic analyses. *Syst. Biol.* 44:548–558.
- Wiens, J. J., and T. W. Reeder. 1997. Phylogeny of the spiny lizards (*Sceloporus*) based on molecular and morphological evidence. *Herpetol. Monogr.* 11:1–101.
- Wilgenbusch, J., and K. de Queiroz. 2000. Phylogenetic relationships among the phrynosomatid sand lizards inferred from mitochondrial DNA sequences generated by heterogeneous evolutionary processes. *Syst. Biol.* 49:592–612.
- Yang, Z., and N. Goldman. 1997. Are big trees indeed easy? *Trends Ecol. Evol.* 12:357.
- Zardoya, R., and A. Meyer. 1998. Complete mitochondrial genome suggests diapsid affinities of turtles. *Proc. Natl. Acad. Sci. U. S. A.* 95:14226–14231.
- Zwickl, D. J., and D. M. Hillis. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Syst. Biol.* 51:588–598.

First submitted 26 November 2002; reviews returned 5 August 2003;

final acceptance 29 March 2004

Associate Editor: Allan Baker