

## Data Partitions and Complex Models in Bayesian Analysis: The Phylogeny of Gymnophthalmid Lizards

TODD A. CASTOE,<sup>1</sup> TIFFANY M. DOAN,<sup>2</sup> AND CHRISTOPHER L. PARKINSON<sup>1</sup>

<sup>1</sup>Department of Biology, University of Central Florida, 4000 Central Florida Boulevard, Orlando, FL 32816-2368, USA;  
E-mail: tcastoe@mail.ucf.edu (T.A.C.), cparkins@pegasus.cc.ucf.edu (C.L.P.)

<sup>2</sup>Biology Department, Vassar College, Box 555, 124 Raymond Avenue, Poughkeepsie, NY 12604-0555, USA;  
E-mail: tiffperu@yahoo.com

**Abstract.**—Phylogenetic studies incorporating multiple loci, and multiple genomes, are becoming increasingly common. Coincident with this trend in genetic sampling, model-based likelihood techniques including Bayesian phylogenetic methods continue to gain popularity. Few studies, however, have examined model fit and sensitivity to such potentially heterogeneous data partitions within combined data analyses using empirical data. Here we investigate the relative model fit and sensitivity of Bayesian phylogenetic methods when alternative site-specific partitions of among-site rate variation (with and without autocorrelated rates) are considered. Our primary goal in choosing a best-fit model was to employ the simplest model that was a good fit to the data while optimizing topology and/or Bayesian posterior probabilities. Thus, we were not interested in complex models that did not practically affect our interpretation of the topology under study. We applied these alternative models to a four-gene data set including one protein-coding nuclear gene (*c-mos*), one protein-coding mitochondrial gene (*ND4*), and two mitochondrial rRNA genes (*12S* and *16S*) for the diverse yet poorly known lizard family *Gymnophthalmidae*. Our results suggest that the best-fit model partitioned among-site rate variation separately among the *c-mos*, *ND4*, and *12S* + *16S* gene regions. We found this model yielded identical topologies to those from analyses based on the *GTR+I+G* model, but significantly changed posterior probability estimates of clade support. This partitioned model also produced more precise (less variable) estimates of posterior probabilities across generations of long Bayesian runs, compared to runs employing a *GTR+I+G* model estimated for the combined data. We use this three-way gamma partitioning in Bayesian analyses to reconstruct a robust phylogenetic hypothesis for the relationships of genera within the lizard family *Gymnophthalmidae*. We then reevaluate the higher-level taxonomic arrangement of the *Gymnophthalmidae*. Based on our findings, we discuss the utility of nontraditional parameters for modeling among-site rate variation and the implications and future directions for complex model building and testing. [Autocorrelated gamma; Bayesian analysis; combining data; *Gymnophthalmidae*; likelihood models; partitioning data; Reptilia; site-specific gamma.]

Incorporating genetic data from multiple genes, often from multiple genomes, is becoming standard in molecular phylogenetics, as is the use of complex model-based likelihood techniques to estimate phylogenetic relationships based on these data. Despite numerous authors advocating the superiority of using multiple loci (especially from multiple genomes) to reconstruct phylogenies (e.g., Pamilo and Nei, 1988; Wu, 1991), few have addressed theoretical and practical effects of modeling sequence evolution simultaneously for different genes (but see examples: Yang, 1996a; Caterino et al., 2001; Pupko et al., 2002; Nylander et al., 2004). Using a single model with a single set of parameters to account for evolution over multiple loci in a combined analysis may fail to accurately portray locus-specific evolutionary patterns. For instance, protein-coding versus rRNA genes may evolve under drastically different constraints because protein-coding genes commonly experience particularly elevated rates of substitution at the third positions of codons. Ribosomal RNA genes, on the other hand, may experience relatively slow rates of compensatory change over regions corresponding to stem-forming secondary structures in the core of the molecule, yet generally more rapid rates in regions corresponding to functionally unconstrained loops and short-range stems (e.g., Dixon and Hillis, 1993; Simon et al., 1994; Muse, 1995; Hickson et al., 1996; Savill et al., 2001). Even among protein-coding genes or rRNA genes, different patterns of substitution rate heterogeneity may result from overall differential rates of evolution or differential functional constraints on partic-

ular regions within a gene (Hickson et al., 1996; Yang, 1996b; Moncalvo et al., 2000). Considering these potential variations in evolutionary rates across sites, it seems logical that models of molecular evolution that account for heterogeneity with regard to among-site rate variation should be employed to reconstruct phylogenetic trees.

The recent shift in phylogenetic methodology towards Bayesian inference of phylogeny has heightened the importance of the use of more realistic evolutionary models. This is important for topological accuracy (e.g., Huelsenbeck, 1995; Huelsenbeck, 1997; Sullivan and Swofford, 2001) as well as accurate estimation of posterior probabilities (e.g., Buckley, 2002; Suzuki et al., 2002; Erixon et al., 2003). In general, it has been shown that likelihood methods are fairly robust to model choice in their estimation of topology (Yang et al., 1994; Posada and Crandall, 2001; Sullivan and Swofford, 2001). A major strength of Bayesian analyses is that posterior probability distributions of trees allow direct interpretation of the likelihood of a particular relationship recovered being true, given the data, the model, and the priors (although the robustness of posterior probabilities has not been thoroughly investigated). However, because the accuracy of posterior probabilities in Bayesian phylogenetic methods relies inherently on the model, models that do not affect the consensus topology may have notable effects on the posterior probability distribution of parameters, and thus on confidence regarding phylogenetic conclusions. Therefore, employing complex

models that more accurately portray DNA evolution should produce less biased posterior probability estimates as long as parameters can be accurately estimated from the data.

The accuracy of posterior probability estimates in Bayesian phylogenetic reconstruction and the factors that may affect this accuracy remain unclear. Many studies suggest Bayesian posterior probabilities appear to be inflated compared to conventional bootstrap support (e.g., Leaché and Reeder, 2002; Cummings et al., 2003; Douady et al., 2003). However, the accuracy of posterior probability support values in terms of both type I and type II error remains unresolved (e.g., Buckley, 2002; Suzuki et al., 2002; Wilcox et al., 2002; Alfaro et al., 2003). Despite this, evidence is accumulating that suggests a direct relationship between accuracy of posterior probabilities and model complexity whereby Bayesian analyses conducted with underparameterized models appear to experience higher error rates compared with parameter rich models (Suzuki et al., 2002; Wilcox et al., 2002; Erixon et al., 2003). However, benefits of constructing and employing more realistic evolutionary models of DNA substitution are challenged by the potential for imprecise and inaccurate parameter estimation (including topology) resulting from overparameterization. Given ever-increasing computational power, in addition to the speed afforded by Bayesian Markov Chain Monte Carlo phylogenetic methods, the need for accurate models and model testing is apparent.

Our primary goal in this paper is to concentrate on evaluation of alternative models that practically affect phylogenetic inference. Specifically, we designed our approaches to make final decisions about best-fitting models based on the effects they had on topology and posterior probability estimates. This is important because although some alternative, relatively parameter-rich models may provide a better fit to the data, they may not result in alternative topologies or significantly different posterior probability estimates. In such cases, our strategy would instead favor a model with fewer parameters that produced essentially the same topology and posterior probability support estimates.

In this study, the genes used to reconstruct phylogenies are diverse and include one protein-coding nuclear gene (*c-mos*), one protein-coding mitochondrial gene (ND4), and two rRNA mitochondrial gene fragments (12S and 16S). We focused on the construction and evaluation of models that utilize alternatively partitioned patterns of among-site rate variation to account for heterogeneous evolution of multiple loci in a combined phylogenetic analysis. Particularly, MrBayes v2.01 allows among-site rate variation ( $\gamma$ ; Yang, 1993) to be partitioned among defined sites (site-specific  $\gamma$ ) as well as allowing the use of an autocorrelated  $\gamma$  parameter to account for local autocorrelation of among-site rates (Kimura, 1985; Schöniger and von Haeseler, 1994; Yang, 1995; Nielsen, 1997). These models allow  $\gamma$  parameter for among-site rate variation to be rescaled across partitions while using a single rate nucleotide substitution rate matrix for the entire data set.

Along with conventional models of sequence evolution (e.g., GTR+I+G), we explore more complex models that partition the among-site rate variation in various ways among loci, in addition to those that employ an additional parameter for autocorrelation of site rate variation. We examine the phylogenetic hypotheses resulting from several alternative partitions of among-site rate variation and discuss their relevance to Bayesian support for clades and support for alternative topological placements of clades.

The taxonomic group examined in this study, lizards of the family Gymnophthalmidae, comprises a large radiation consisting of approximately 34 genera and 180 species occurring throughout South America with relatively few species in Middle America (Pellegrino et al., 2001; Doan, 2003a). The family is composed of small to medium lizards that occur in a variety of habitats and occupy a wide range of niches. This lizard group has been poorly studied with many species unknown beyond their original descriptions.

Relationships of genera within the family Gymnophthalmidae are poorly understood. The most comprehensive and contemporary revision of the supergeneric classification of the family Gymnophthalmidae was made by Pellegrino et al. (2001). They reconstructed a phylogeny of 50 species in 24 genera (recently reduced from 26 by Doan, 2003a) using five genes (two nuclear and three mitochondrial). Based on their reconstruction they erected four subfamilies and four tribes. The subfamily Alopoglossinae, consists solely of *Alopoglossus*. The subfamily Gymnophthalminae contains 13 genera, divided into two tribes, the Heterodactylini (5 genera) and the Gymnophthalmini (8 genera). The subfamily Rhachisaurinae is monotypic, consisting of *Rhachisaurus brachylepis*, a new genus separated from *Anotosaura*. The final subfamily, the Cercosaurinae, consists of 20 genera divided into tribe Cercosaurini (14 genera) and tribe Eupleopini (6 genera).

Harris (2003) used *c-mos* sequences to reconstruct a phylogeny of the Squamata with concentrated taxon sampling in the Gymnophthalmidae. He primarily used Pellegrino et al.'s (2001) sequences but added a new sequence of *Proctoporus bolivianus*. Harris's (2003) reconstruction differed from Pellegrino et al.'s in the placement of *Ptychoglossus*, *Bachia*, *Arthrosaura*, and several smaller scale relationships. In addition, a teiid genus, *Tupinambis*, was nested within the Gymnophthalmidae as the sister to *Ptychoglossus* and *Alopoglossus* (although this relationship received bootstrap and posterior probability support below 50%).

Missing from Pellegrino et al.'s (2001) study were 10 genera. Whereas Pellegrino et al. sampled all genera of Alopoglossinae, Rhachisaurinae, and Gymnophthalmini, they lacked *Stenolepis* from the Heterodactylini, *Amapasaurus* from the Eupleopini, and eight genera from the Cercosaurini. The limited taxon sampling of the Cercosaurini renders conclusions about that tribe problematic (see Hillis, 1998) because only half of the genera and 18 of the approximately 121 species were sampled (14.9%).

In addition to limited taxon sampling, the separate gene partitions were often in conflict with regards to the positions and relationships of many key taxa (Pellegrino et al., 2001). Such conflicts put into doubt the subfamilial and/or tribal placement of genera such as *Ptychoglossus*, *Rhachisaurus*, *Bachia*, and *Neusticurus*. Our study addresses some of the problems suggested in the combined analysis of Pellegrino et al. (2001) and fills in some significant gaps in taxon sampling. Concentrating on the Cercosaurini, we add 19 new individuals, including 12 new species and one new genus, as well as an additional individual of *Ptychoglossus brevifrontalis* (a total of 73 additional sequences). With this greater taxon sampling we clarify the relationships and classification within family Gymnophthalmidae. In addition to adding more taxa, we utilize a Bayesian approach to the phylogeny to complement the standard parsimony and likelihood methods used by Pellegrino et al. (2001). We synthesize this information, emphasizing overall phylogenetic evidence, and propose an alternative hypothesis for the intergeneric relationships and taxonomy within the family.

The objectives of our study included (1) evaluating the effects of partitioning among-site rate variation and among-site rate autocorrelation parameters on phylogenetic topology and posterior probabilities for relationships, (2) developing a robust strategy for choosing the best-fit model for among-site rate variation considering practical effects on topology and posterior probability support, (3) identifying the most likely and robust phylogenetic hypothesis for relationships among gymnophthalmid lizards, and (4) reevaluating the supergeneric classification of the family based on our best estimate of gymnophthalmid phylogeny.

## METHODS

### *DNA Sequences Used*

A significant subset of the sequences used in this study is from Pellegrino et al. (2001) and Doan and Castoe (2003). Additional sequences of gymnophthalmid lizards were added to this dataset. Of the five genes used by Pellegrino et al. (2001), we chose to use and expand upon four: mitochondrial NADH dehydrogenase subunit 4 (ND4), mitochondrial small subunit rRNA gene (12S), mitochondrial large subunit rRNA gene (16S), and the nuclear oocyte maturation factor gene (*c-mos*). The nuclear small subunit rRNA gene (18S) used by Pellegrino et al. (2001) was omitted from our study for two reasons: (1) low phylogenetic signal apparent from the Pellegrino et al. (2001) study, and (2) the nuclear gene for 18S occurs in hundreds or thousands of copies per nuclear genome (e.g., humans: International Human Genome Sequencing Consortium, 2001; *Xenopus*: Pardue, 1974; Long and Dawid, 1980). Using sequences of this multicopy gene to resolve relationships principally among species and genera within a family may increase the potential for recovery of misleading phylogenetic estimates based on incomplete gene conversion among alleles at different loci or differential fixation of alleles

among loci (Gasser et al., 1998; Gonzalez and Sylvester, 2001).

Laboratory methods for obtaining novel sequences used in this study are as follows. Where possible, two individuals of each taxon from distant sampling localities were added to the data set. Genomic DNA was isolated from tissue samples (liver or skin preserved in ethanol) using the Qiagen DNeasy extraction kit and protocol (Qiagen Inc., Hilden, Germany). The mitochondrial ND4 gene was amplified via polymerase chain reaction (PCR) using the primers ND4 and LEU as described in Arévalo et al. (1994). Mitochondrial ribosomal small and large subunit genes (12S and 16S) were amplified as described in Parkinson et al. (1997, 1999). The nuclear *c-mos* gene was amplified with primers G73 and G74 as described in Saint et al. (1998) and Pellegrino et al. (2001). Positive PCR products were excised from agarose electrophoretic gels and purified using the GeneCleanIII kit (BIO101). Purified PCR products were sequenced in both directions with the amplification primers (and for ND4, an additional internal primer HIS, Arévalo et al., 1994). Samples that could not be sufficiently sequenced directly were cloned using the Topo TA cloning kit (Invitrogen) according to the manufacturer's protocols. Plasmids were isolated from multiple clones per individual using the Qiaquick spin miniprep kit (Qiagen). Plasmids were sequenced using M13 primers (provided by Topo TA kit, Invitrogen) and, in some cases, the internal HIS primer for ND4. Purified PCR products and plasmids were sequenced using the CEQ D Dye Terminator Cycle Sequencing (DTCS) Quick Start Kit (Beckman-Coulter) and run on a Beckman CEQ2000 automated sequencer according to the manufacturers' protocols. Raw sequence chromatographs for sequences generated in this study were edited using Sequencher 3.1 (Gene Codes Corp.). In cases where gene fragments were cloned, all sequences from a single individual were edited together. Novel sequences were deposited in GenBank. The GenBank accession number for each gene sequence used in this study (including novel sequences) are given in Appendix 1.

### *Sequence Homology and Alignment*

Multiple sequence alignment was performed using ClustalW (Thompson et al., 1994). Initial alignments were conducted with a gap opening penalty of 10, a gap extension penalty of 1, and a transition weight of 0.5. For rRNA genes (12S and 16S), alternative multiple alignments were examined with gap opening and gap extension penalties ranging from 10 and 10 (respectively) to 1 and 1, including varying ratios in this range. Initial alignments for protein-coding genes (ND4 and *c-mos*) were rechecked based on the homology of their translated amino acid sequence using GeneDoc (Nicholas and Nicholas, 1997). The ND4 alignment was unambiguous and not edited manually and the *c-mos* alignment was slightly manually modified to maximize amino acid similarity over a short indel region within the alignment. Alternative automated alignments (from ClustalW) for rRNA genes (12S and 16S) were compared,

along with estimates of secondary structures (Gutell, 1994; Gutell et al., 1994; Titus and Frost, 1996), to evaluate evidence for positional homology. Positions in rRNA genes where alignment was ambiguous were excluded. To minimize the effects of missing data resulting from incomplete sequences, all gene alignments were truncated at the 5' and 3' ends. The final alignment of all concatenated genes, including positions excluded from phylogenetic analyses, is available as supplemental data at <http://biology.ucf.edu/~clp/Lab/Lab.htm>.

#### *Phylogeny Estimation Using Maximum Parsimony*

We inferred phylogenies based on the maximum parsimony (MP) criterion in PAUP\* v4.0b10 (Swofford, 2002) and Bayesian (Markov Chain Monte Carlo, MCMC) analysis in MrBayes v2.01 (Huelsenbeck and Ronquist, 2001). Phylogenetic inference was conducted, hierarchically, in three steps: (1) all genes individually, (2) intermediate partitions including both rRNA genes (12S + 16S) and all mitochondrial genes (mtgenes), and (3) the combined concatenated data set including all four genes. For MP analyses of independent genes and intermediate partitions, we conducted equally weighted parsimony searches using the heuristic strategy with 200 random taxon addition sequence replicates. Settings for MP analyses were tree bisection-reconnection branch swapping, steepest descent off, and MULTREES option on (Swofford, 2002). For all individual genes and intermediate partitions (rRNA and mtgenes), we assessed support for clades using 200 nonparametric bootstrap pseudoreplicates (Felsenstein, 1985) with 20 random taxon addition sequence replicates implemented with PAUP\*. For the combined MP analysis of all genes, we searched for trees using equally weighted parsimony heuristic searches with 1000 random taxon addition sequence replicates and assessed clade support with 200 bootstrap pseudoreplicates with 200 random addition sequence replicates per bootstrap pseudoreplicate. We consider relationships that are supported by at least 70% bootstrap to be significantly resolved (Hillis and Bull, 1993).

#### *Bayesian Phylogeny Estimation*

ModelTest version 3.0 (Posada and Crandall, 1998) was used to infer the best-fit model of evolution for each gene data set (individual genes, intermediate partitions, and total combined data) based on hierarchical log-likelihood ratio tests comparing successively complex models (Huelsenbeck and Crandall, 1997; Posada and Crandall, 2001).

All MCMC phylogenetic reconstructions were conducted in MrBayes v2.01 (Huelsenbeck and Ronquist, 2001) with vague priors (as per the program's defaults) and model parameters estimated as part of the analyses. Three heated chains and a single cold chain were used in all MCMC analyses and runs were initiated with random trees, as per the program's defaults. Trees were sampled every 100 generations and majority rule consensus phylograms and posterior probabilities for nodes were assembled from all post burn-in sampled trees. Phylo-

genetic reconstructions for all data partitions were estimated using three independent runs to confirm that stationarity (or global optimality) was reached and that independent runs converged on similar stationary parameter estimates. Each of these data partition runs was conducted with a total of 1.4 million generations, 400,000 of which were discarded as burn-in, yielding 1 million post burn-in generations.

Each MCMC run for all individual gene and intermediate data partitions employed the model selected by ModelTest for that partition, or the nearest model to that model that could be implemented in MrBayes. The total combined data set was subjected to MCMC analyses under multiple alternative evolutionary models which differed in the way they parameterized among-site rate variation.

The most complex (parameter rich) model that ModelTest v3.0 can evaluate is a general time reversible (GTR; Tavaré, 1986) model with an estimated proportion of invariant sites (I) and gamma distributed among-site rate variation (G; Yang, 1993). MrBayes v2.01 is capable of employing more complex models than this GTR+I+G model. MrBayes allows among-site rate variation to be partitioned among user defined sites (site-specific gamma; SSG) as well as allowing the use of an autocorrelated gamma (A; e.g., Yang, 1995; Penny et al., 2001; Huelsenbeck, 2002) to account for autocorrelation of among-site rates. These two modifications of among-site rate variation may be used independently as well as simultaneously in a given model in MrBayes. In addition to the GTR+I+G model, we conducted combined data MCMC analyses with alternative models that partitioned gamma with (SAG) and without (SSG) accounting for autocorrelation (A) of rates. These two alternative ways to estimate among-site rate variation were invoked by the commands "rates = ssadgamma" and "rates = ssgamma" (respectively) in the MrBayes 2.01 command block. All models applied a single common GTR substitution rate matrix across all data and differed only in the way they modeled and partitioned among-site rates according to the following a priori partitions: GTR + autocorrelated gamma (GTR+AG); protein-coding genes versus rRNA genes (PR-SSG and PR-SAG); nuclear versus mitochondrial genes (NM-SSG and NM-SAG); c-mos versus ND4 versus rRNA genes (CNR-SSG and CNR-SAG); all genes partitioned independently (4gene-SSG and 4gene-SAG). Table 1 provides a summary of the parametric content of each of these models.

Choosing among these models to identify the best model of evolution on which to base phylogenetic and taxonomic decisions was approached in several ways. Our goal was to find the model of evolution which best fit the data yet contained the fewest total parameters (the best-fit model). Specifically, our major criteria for identification of the simplest best-fit model included the demonstration of clear improvements of likelihood estimates under that model, along with a practical effect on topology and/or posterior probability support for clades. Therefore, we were not interested in more complex models that did not estimate a different topology

TABLE 1. Parametric composition of models tested in Bayesian MCMC analyses of the combined data.

Model name	Nucleotide substitution matrix	Model parameters in addition to GTR matrix	Gamma parameter	Gamma autocorrelation parameter	Site (= partition)-specific gamma parameters
GTR+I+G	GTR	2	+	–	–
GTR+AG	GTR	2	+	+	–
NM-SSG	GTR	3	+	–	2
PR-SSG	GTR	3	+	–	2
NM-SAG	GTR	4	+	+	2
PR-SAG	GTR	4	+	+	2
CNR-SSG	GTR	5	+	–	3
CNR-SAG	GTR	6	+	+	3
4gene-SSG	GTR	6	+	–	4
4gene-SAG	GTR	7	+	+	4

or have significant effects on posterior probability estimates. Once a tentative model was chosen, this model was rigorously tested for overparameterization and unreliability (which would suggest it was not a candidate for the best-fit model).

We examined the burn-in plots of likelihoods for MCMC chains for each model to determine the rate of ascent to an apparent stationary plateau, in addition to the degree of overlap between models and superiority (based on chain likelihood values) of models, relative to the number of parameters they employed. To examine the relative improvement in likelihood scores with respect to model complexity, we compared the 95% credibility interval (CI) of MCMC chain likelihood scores between models. To calculate the 95% CI, we ranked all post burn-in tree estimates by ln likelihood and included the most likely 95% (Felsenstein, 1968; Huelsenbeck et al., 2002). Once a tentative best-fit was identified, we evaluated parameter burn-in plots of these models for evidence of identifiability of parameters by checking for commonality in parameter estimates among runs. We also examined the sensitivity of posterior probability values to model complexity using Wilcoxon signed rank tests implemented with Statistica (StatSoft, 1993) to test for significant changes in posterior probability estimates between the chosen model and those which were proximal alternative best-fit models. For interpretation of phylogenetic inferences, we consider posterior probability values over 95% to be well resolved.

In addition to the three independent MCMC runs (1.4 million generations each) conducted for each model, we conducted a single MCMC run for an extended number of generations (33 million generations) for the two main alternative best-fit models (GTR+I+G and CNR-SSG) for the combined data set. For each long MCMC run, the posterior probabilities for clades were monitored in intervals of two million generations to examine any trends and the overall precision associated with posterior probability estimates through generations of extended MCMC runs. Additionally, posterior probabilities of clades estimated from these long MCMC runs were compared to those estimated from the initial three MCMC runs per model (run for 1.4 million generations) to examine the effect that MCMC analysis strategy (multiple short runs versus single long run) has on estimates of posterior probabilities.

Posterior probabilities estimated from these long runs were also used to re-test for significant changes in posterior probability estimates derived from analyses under alternative models (as described above).

## RESULTS

A total of 1810 characters were included in the analysis (c-mos 408 bp; ND4 623 bp; 12S 331 bp; 16S 448 bp). Details of optimal trees selected by maximum parsimony and best-fit models of evolution selected by ModelTest (Posada and Crandall, 1998) are presented in Table 2. After preliminary phylogenetic reconstructions, we identified several apparent problems with the Pellegrino et al. (2001) data set, including switching of taxon names and apparent contamination, which we rectified prior to final analyses. These are discussed in detail in Appendix 2.

### *Parsimony Phylogenetic Reconstruction*

The total evidence (all four genes) equally weighted parsimony reconstruction resulted in two most parsimonious trees of 6600 steps with 769 parsimony-informative characters and CI = 0.228, RI = 0.543, RC = 0.124, HI = 0.772 (Fig. 1). Six major clades were recovered, each with high bootstrap support (70% to 100%) and differing from the reconstruction of Pellegrino et al. (2001). Whereas the earliest split within the Gymnophthalmidae in Pellegrino et al. (2001) was the divergence of a clade composed of the three *Alopoglossus* species from all others, we recovered a clade of *Alopoglossus* spp. and *Ptychoglossus brevifrontalis*. As explained in Appendix 2, an apparent taxon name error in the 12S and 16S data sets presumably resulted in the erroneous placement of *Ptychoglossus* in the Cercosaurinae. In addition to the *Alopoglossus* + *Ptychoglossus* clade, five other major clades were present, each of which had bootstrap support of 70% or greater. Relationships among these clades did not receive high bootstrap support. The Heterodactylini was sister to the Gymnophthalmini, supporting a monophyletic Gymnophthalminae with high bootstrap support (99%). The other three clades included a clade of *Bachia* + *Rhachisaurus*, the Ecleopini, and the Cercosaurini. Monophyly of the Cercosaurinae did not receive strong bootstrap support.

TABLE 2. Statistics for datasets used, including results from MP searches and suggested model from hierarchical ln likelihood ratio test (hLRT) criterion from ModelTest.

	c-mos	ND4	12S	16S <sup>a</sup>	All rRNA	All mt	All protein	Total
Number of characters	408	623	331	448	779	1402	1031	1810
Parsimony-informative	173	363	112	121	233	596	536	769
Number of trees	666	30	5195	>120,000	5065	6	4	2
Optimal tree score	566	4365	749	739	1547	5990	4976	6600
CI	0.528	0.177	0.290	0.296	0.282	0.202	0.215	0.228
HI	0.472	0.823	0.710	0.704	0.718	0.798	0.785	0.772
hLRT-selected model	K80+G	GTR+I+G	TrN+I+G	TrN+I+G	HKY+I+G	TVM+I+G	TVM+I+G	GTR+I+G
Proportion invariable sites	—	0.321	0.421	0.551	0.535	0.461	0.331	0.432
Gamma parameter	0.595	0.506	0.543	0.511	0.657	0.564	0.519	0.553
Ti:Tv ratio	2.67	—	—	—	2.651	—	—	—
Rate matrix								
r(A-C)	—	0.301	1	1	—	0.438	0.411	0.553
r(A-G)	—	6.907	12.086	3.023	—	3.531	4.370	4.101
r(A-T)	—	0.648	1	1	—	0.476	0.543	0.530
r(C-G)	—	0.429	1	1	—	0.179	0.630	0.384
r(C-T)	—	4.648	5.352	7.150	—	3.531	4.370	3.608
r(G-T)	—	1	1	1	—	1	1	1

<sup>a</sup>We were unable to complete the 16S run due to the enormous computational time required to store and evaluate the high number of optimal trees. To make searching this data set feasible, we ran a heuristic search with 8000 random addition sequence replicates saving no more than 50 trees at each step.

#### Mitochondrial Gene MCMC Analyses

Based on hierarchical log likelihood ratio tests (hLRTs) of successively complex models of sequence evolution, ModelTest indicated the best-fit model for the combined mitochondrial dataset was the TVM+I+G (Table 2). This model of evolution, characterized by a five-parameter nucleotide substitution rate matrix, is not currently available in MrBayes. Instead, the next best-fitting parameter rich model, which employs a GTR six-parameter nucleotide substitution rate matrix, was employed with proportion of invariant sites (I) and gamma distributed among-site rate variation (G). Parameter estimates derived from the combination of all post burn-in estimates from the three independent MCMC runs are summarized in Table 3. All three runs reached apparent stationarity (in estimates of substitution model parameters, as well as chain likelihood scores) prior to 50,000 generations, well before the conservative burn-in period of 400,000 generations.

The all mitochondrial data partition MCMC reconstruction (Fig. 2a) contrasts with the Pellegrino et al. (2001) reconstruction in the relative phylogenetic place-

ment of major clades deep in the phylogeny, but posterior probability support for some the relationships was not high. As in the parsimony reconstruction described above, *Alopoglossus* and *Ptychoglossus* form a clade sister to the remaining gymnophthalmids. The next node splits the Eupleopini from the remainder of the taxa (but with low posterior probability support). Of the three MCMC runs, one differed slightly with regard to the structure of the remainder of the tree. Examination of this difference among runs revealed that the difference between the majority rule topologies from post burn-in MCMC trees resulted from an approximately 1% difference between runs in the posterior probability density supporting one relationship over another (both of which received posterior probabilities below 50%). Figure 2a depicts the consensus of those three runs that collapses nodes in conflict, creating a polytomy of *Rhachisaurus*, *Bachia*, the Gymnophthalminae, and the Cercosaurini minus *Bachia*. Even with the differences among the reconstructions, the lack of monophyly of the Cercosaurinae differs from Pellegrino et al. (2001; their Fig. 4) and our parsimony reconstruction (Fig. 1).

TABLE 3. Parameter estimates for all mitochondrial gene and c-mos.

	All mt genes—All runs	c-mos—All runs
ln likelihood	-26274.1 (-26295.6 to -26254.2)	-3616.0 (-3634.3 to -3599.0)
pi(A)	0.397 (0.380-0.416)	0.259 (0.0.230-0.289)
pi(C)	0.283 (0.269-0.297)	0.265 (0.236-0.294)
pi(G)	0.077 (0.070-0.084)	0.243 (0.215-0.272)
pi(T)	0.243 (0.230-0.256)	0.233 (0.205-0.262)
r(A-C)	0.406 (0.310-0.524)	—
r(A-G)	4.236 (3.497-5.091)	—
r(A-T)	0.470 (0.351-0.604)	—
r(C-G)	0.201 (0.124-0.300)	—
r(C-T)	3.641 (2.910-4.459)	—
r(G-T)	1	—
Tv:Ti ratio	—	5.432 (4.490-6.534)
Gamma parameter	0.503 (0.462-0.547)	0.644 (0.517-0.803)
Proportion of invariable sites	0.419 (0.387-0.450)	—

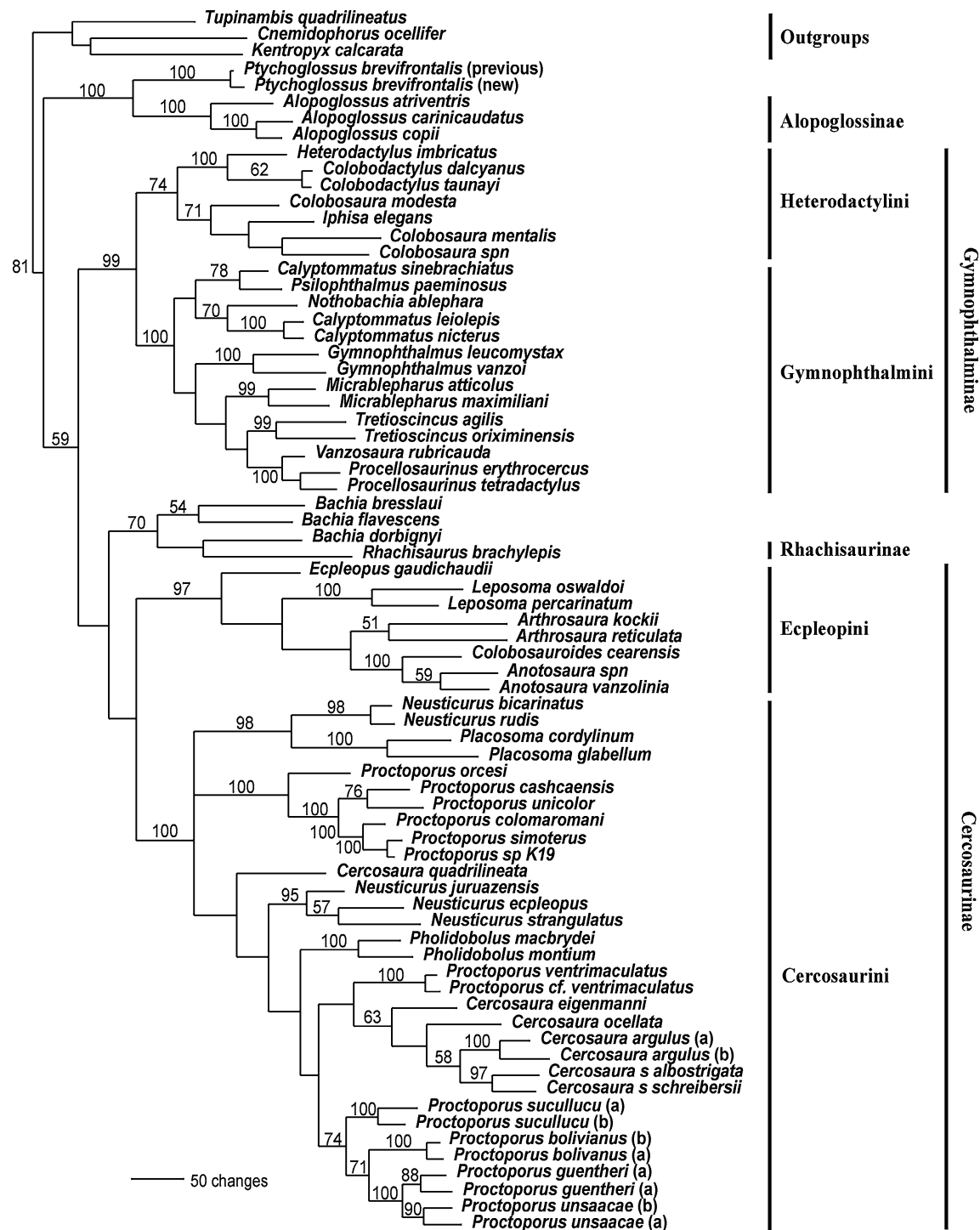


FIGURE 1. Strict consensus phylogram of two most parsimonious trees based on the equally weighted maximum parsimony search including all four genes (*c-mos*, *ND4*, *12S*, and *16S*). Labels (a) and (b) indicate individuals of a species (see Appendix 1). For reference, labels on the right side represent the taxonomy presented by Pellegrino et al. (2001). Taxa that are not labeled have relationships that do not agree with that former taxonomy.

#### *C-mos* (Nuclear Gene) MCMC Analyses

Based on hLRTs of successively complex models of sequence evolution, ModelTest indicated the best-fit model for the combined mitochondrial dataset was the

K80+G model (Table 2). Parameter estimates derived from the combination of all post burn-in estimates from the three independent MCMC runs, using a K80+G model, are summarized in Table 3. All three runs reached

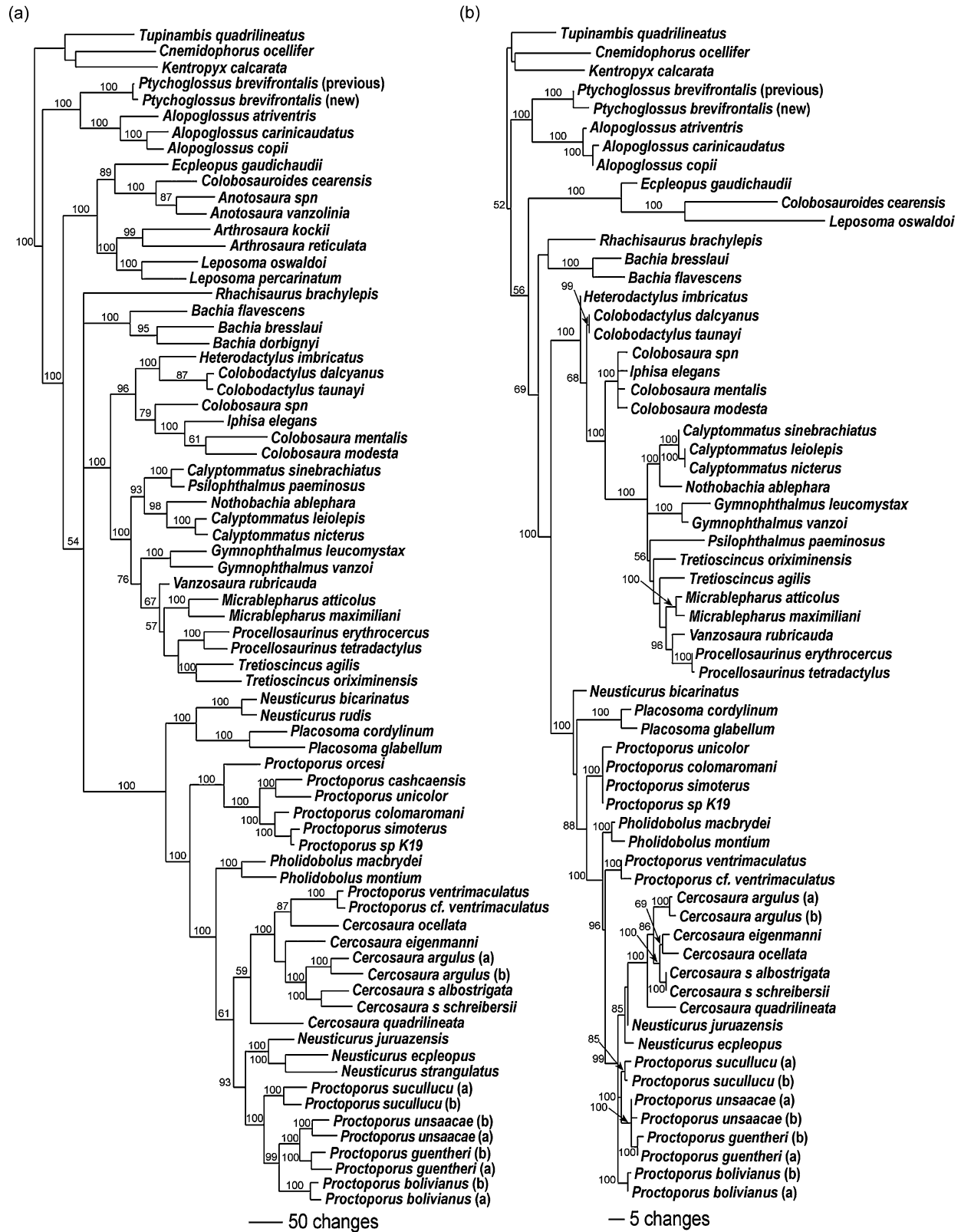


FIGURE 2. Bayesian phylogenetic trees for the independent nuclear and mitochondrial data partitions. Labels (a) and (b) indicate individuals of a species (see Appendix 1). (a) Majority rule phylogram and posterior probabilities resulting from Bayesian analysis of all three mitochondrial genes combined (ND4, 12S, and 16S) based on a combined 3 million post burn-in generations under the GTR+I+G model of evolution. (b) Majority rule phylogram and posterior probabilities resulting from Bayesian analysis of nuclear c-mos gene data based on a combined 3 million post burn-in generations under the K80+G model of evolution.



apparent stationarity (in estimates of substitution model parameters, as well as chain likelihood scores) prior to 50,000 generations.

The nuclear *c-mos* reconstruction (Fig. 2b) differs from that of Pellegrino et al. (2001), Harris (2003), in our parsimony reconstruction (Fig. 1), and our mitochondrial reconstruction (Fig. 2a). As with Harris (2003), in our parsimony reconstruction, and our mitochondrial DNA reconstruction, *Alopoglossus* and *Ptychoglossus* form a basal clade. Similar to our parsimony reconstruction, four additional major clades are formed, each with high posterior probability support for clade monophyly, but low support of the relationships among the clades. The Gymnophthalminae forms a monophyletic group with strong posterior probability support. The Cercosaurinae is not monophyletic because there is strong support for the Ecleopini being only distantly related to the Cercosaurini. Additionally, as in the parsimony reconstruction, *Bachia* and *Rhachisaurus* form a clade. As in Pellegrino et al.'s (2001) maximum likelihood reconstruction, but differing from our parsimony reconstruction, tribe Heterodactylini is not monophyletic but is paraphyletic with respect to the Gymnophthalmini.

#### Combined MCMC Analyses

Based on the hLRT criterion for model selection, ModelTest chose the GTR+I+G model of nucleotide substitution for the combined data set (Table 2). Burn-in plots of likelihood scores of MCMC chains conducted with this model and alternative models are shown in Figure 3 and mean stationary values (with 95% CI) across mod-

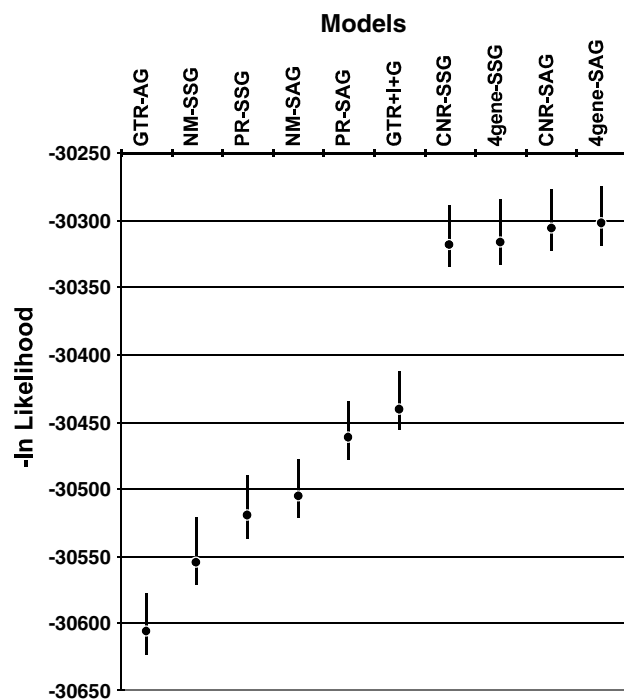


FIGURE 4. The mean and 95% credibility interval for post burn-in  $\ln$  likelihood scores of MCMC chains based on alternative models of evolution. See text for descriptions of alternative models.

els are compared in Figure 4. A more detailed plot of the ascent of likelihood scores of chains toward stationarity for each model is shown in Figure 5. Although we only show burn-in plots for chains from one of three

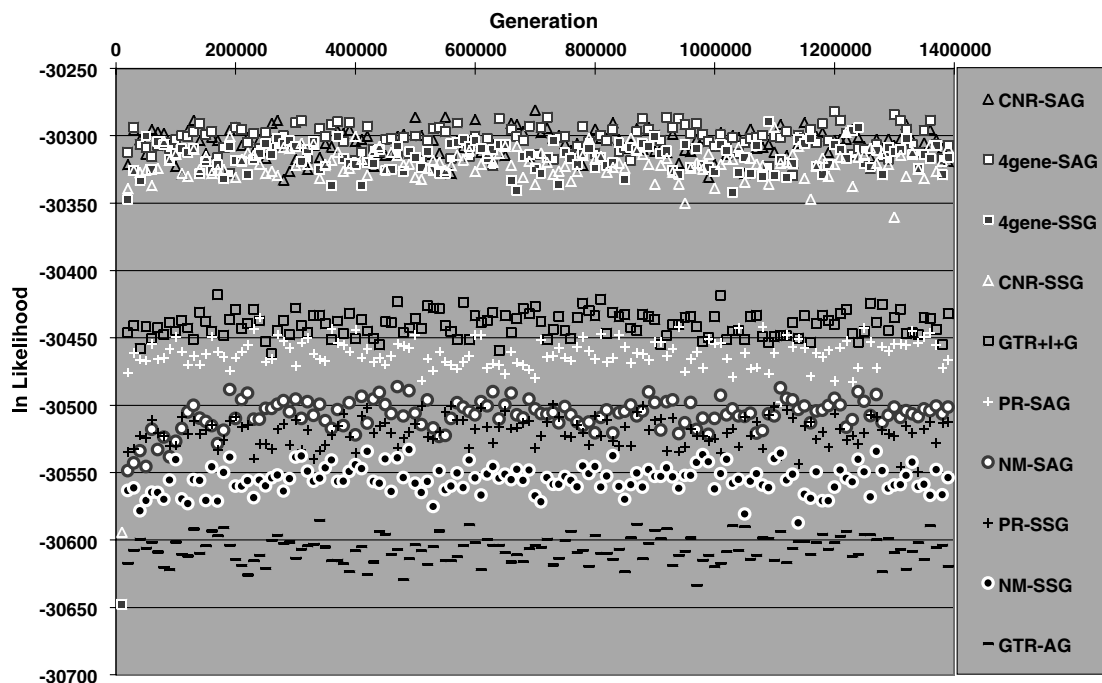


FIGURE 3. The  $\ln$  likelihood scores of MCMC chains based on alternative models of evolution, sampled in 10,000 generation intervals for clarity of presentation. See text for descriptions of alternative models.

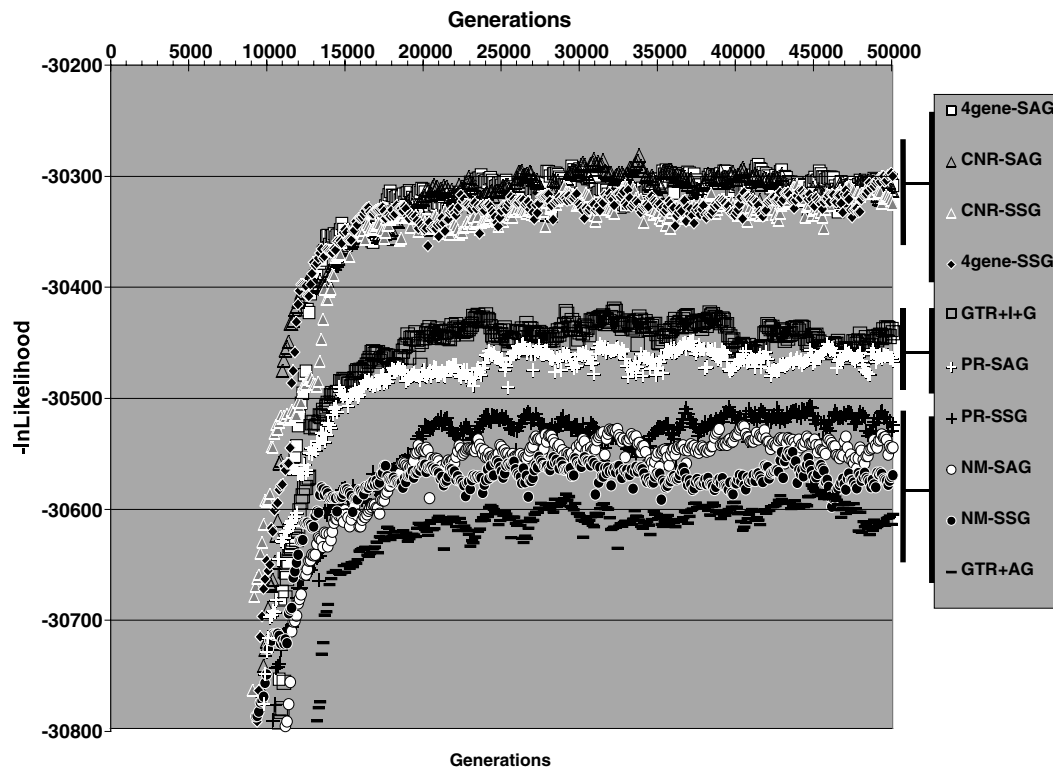


FIGURE 5. The  $\ln$  likelihood scores of MCMC chains based on alternative models of evolution, focusing on the period below 50,000 generations, sampled every 100 generations. See text for descriptions of alternative models.

individual MCMC analyses for each model, no single run (under a particular model) was noticeably different with regard to burn-in time, parameter estimate mean, or CI at stationarity.

Based on the post-burn-in plateau of chain likelihood values observed in Figures 3 and 4, the GTR+I+G model appears to out-perform models that partition among-site rate variation between either nuclear versus mitochondrial genes (NM-SSG and NM-SAG) or between protein-coding versus ribosomal RNA genes (PR-SSG and PR-SAG). The GTR+AG model resulted in chain likelihood scores that were markedly lower than those estimated under the GTR+I+G model. Two groups of models that partition among-site rate variation into either three or four classes appeared to result in clear improvements in the likelihood scores of stationary chains when compared to the GTR+I+G model: models that partitioned among-site rate variation among *c-mos* (nuclear protein-coding) versus *ND4* (mitochondrial protein-coding) versus ribosomal RNA genes (CNR-SSG and CNR-SAG; three site partitions) and those that partitioned rate variation among all individual genes (4gene-SSG and 4gene-SAG; four site partitions). Within this group of models with either three or four partitions of among-site rate variation (with and without autocorrelated gamma), no single model clearly outperformed any other based on estimates of stationary chain likelihood scores (Fig. 3). From Figure 5 we observe that all models, including those with three or four partitions for among-site rate

variation, achieve stationarity rapidly by approximately 30,000 generations (although we conservatively discarded trees prior to 400,000 generations as burn-in).

Consensus topologies estimated from post-burn-in generations were identical among multiple independent runs under a particular model (Fig. 6). We found a general correlation between topology and model fit (inferred based on relative values of stationary chain likelihood scores), whereby the two models with the lowest range of  $\ln$  likelihood scores (mean  $\ln$  likelihood less than  $-30,550$ ) for chains produced slightly different topologies compared with all models resulting in chains with higher  $\ln$  likelihoods (mean  $\ln$  likelihood less than  $-30,550$ ). Analyses of the combined data employing all models except NM-SSG and GTR+AG recovered the identical topology. The analyses under the NM-SSG and GTR+AG models recovered a topology identical to the others except for a swap in the relative branching order with respect to two clades (*Rhachisaurus* + *Gymnophthalminae* and *Epleopini*; neither rearrangement received high posterior probability support), in addition to a modification affecting the phylogenetic position of *Proctoporus ventrimaculatus* + *P. cf. ventrimaculatus*.

Based on our a priori criteria for initial identification of the preferred evolutionary model as that which contained the fewest number of parameters while demonstrating a clear optimization of overall chain likelihood, we chose the CNR-SSG model. To examine the effect of

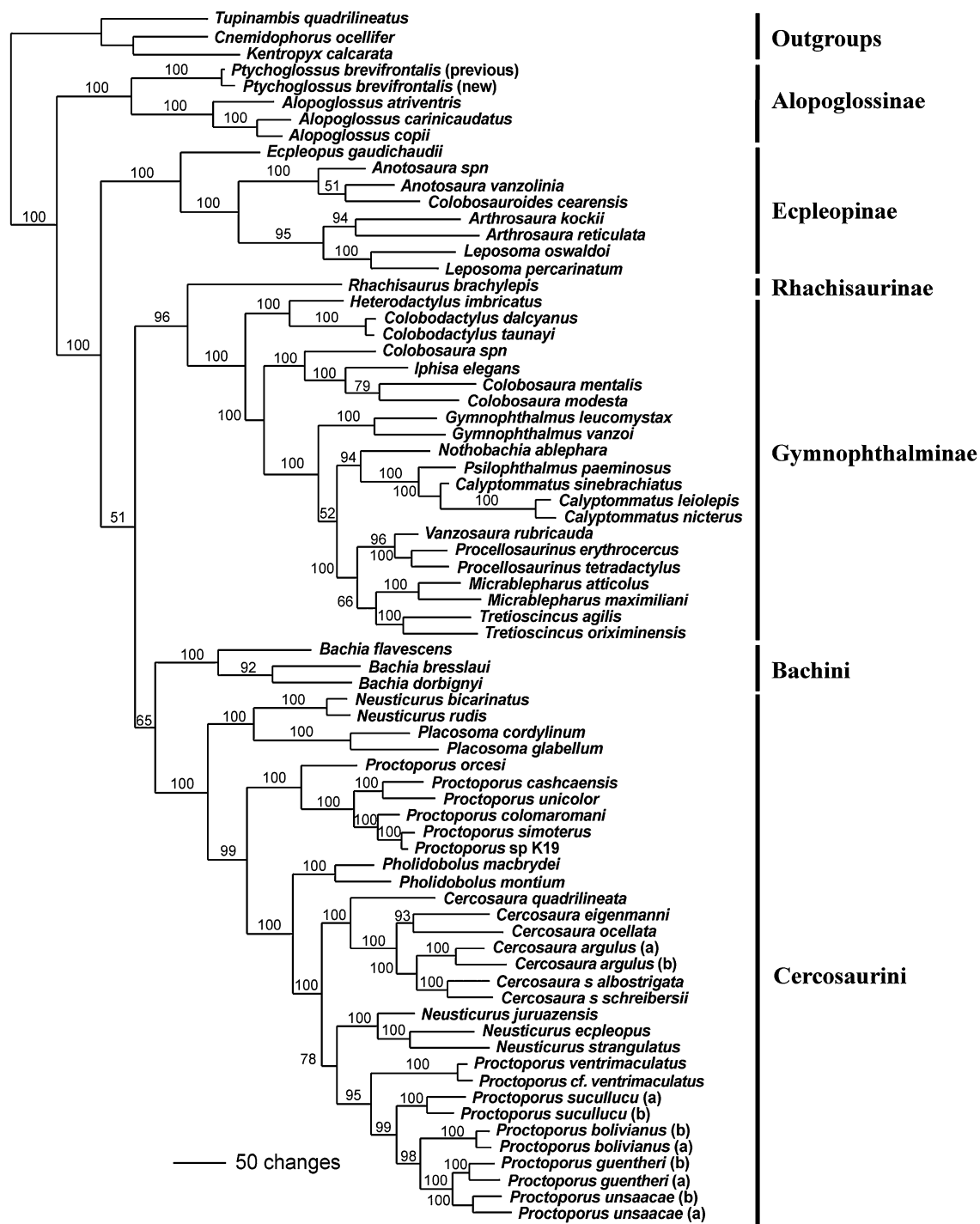


FIGURE 6. Bayesian phylogenetic tree and posterior probabilities for clades based on the combined, four-gene data set analyzed under the CNR-SSG model. Tree is based on the combination of all post burn-in generations resulting from three independent runs of the model, for a combined total of 3 million post-burn-in generations. Labels (a) and (b) indicate individuals of a species (see Appendix 1). This work's new proposed phylogenetic classification depicted by labeled bars on the right.

model choice on the cumulative posterior probabilities for clades, we tested for significant changes in posterior probabilities between the CNR-SSG model and all other models that were found to be as good or better

than the GTR+I+G model (including GTR+I+G, CNR-SAG, 4gene-SSG, and 4gene-SAG) with Wilcoxon signed rank tests. For these, posterior probabilities for matched nodes were pairs for comparison. Tests comparing

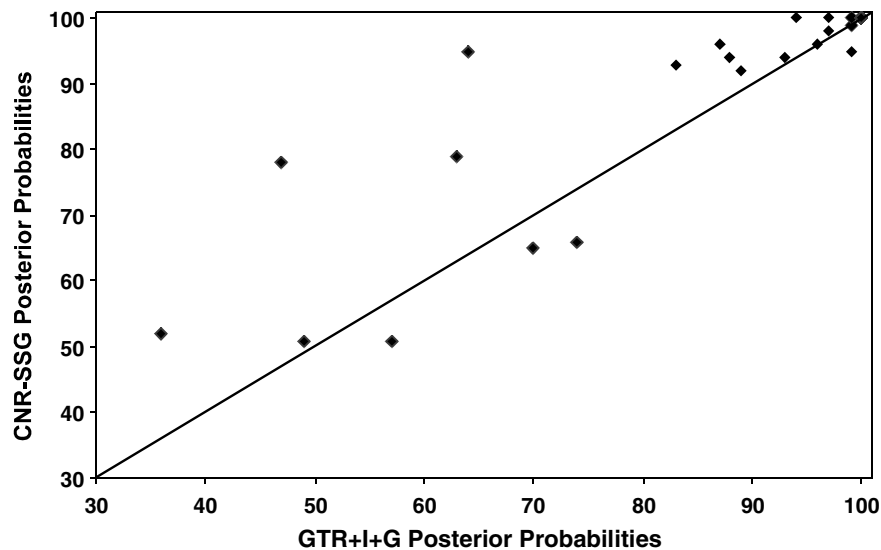


FIGURE 7. Plot of the posterior probabilities derived from the GTR+I+G MCMC analyses (all three runs combined) versus the posterior probabilities derived from the CNR-SSG model (all three runs combined). For comparison, a 1:1 line is plotted on the same axis.

the overall change in posterior probabilities between the CNR-SSG model and other models with three or four gamma partitions (with and without autocorrelated gamma) were not significant. However, the GTR+I+G model was found to produce, overall, significantly lower estimates for posterior probabilities of clades than the preferred model (CNR-SSG;  $z = 2.173$ ,  $P = 0.029$ ). This trend is demonstrated by the relationship between posterior probabilities from the GTR+I+G model and the CNR-SSG model, plotted against one another in Figure 7. Overall, a majority of nodes plotted in this figure fall above the 1:1 line, indicating higher nodal support resulting from the CNR-SSG model. It is also important to note, however, that several nodes did decrease in posterior probability support under the CNR-SSG model.

As described above, burn-in plots of ln likelihood of MCMC chain scores from all independent MCMC runs under the CNR-SSG model are essentially identical with a rapid and direct approach to a common stationary plateau (not shown). To investigate burn-in and common estimates at stationarity for the parameters of the independent (1.4 million generation) CNR-SSG model runs, burn-in plots of the reversible rate of A-G and A-T substitutions as well as the gamma parameter and site-specific partition rate parameters are shown in Figure 8. Similar to burn-in plots of likelihood tree scores (Fig. 5), all parameters appear to approach stationarity rapidly (in less than 50,000 generations) and oscillate around a common stationary value (across independent runs). Because all three (1.4 million generation) independent runs of our preferred model (CNR-SSG) appear to reach common stationary estimates of parameters, produce identical topologies, and nearly identical posterior probability estimates, hereafter we report only results based on the combination of all 3 million post burn-in generations pooled from the three independent runs of MCMC analyses using the CNR-SSG model.

Parameter values, with 95% credibility intervals, resulting from MCMC CNR-SSG model analyses are given in Table 4.

Posterior probability estimates derived from post burn-in generations from the single long MCMC run (33 million generations) of the GTR+I+G and CNR-SSG models were very similar to estimates based on the combination of the three shorter (1.4 million generation) runs. Considering only clades supported with less than 100% posterior probability support, the long MCMC run of the GTR+I+G model produced estimates that were, on average, 1.05% different from the short run estimates, compared with 0.40% for the CNR-SSG model. Given the almost identical posterior probability estimates (within 1%) derived from the single long MCMC run under the CNR-SSG model as compared with those previously estimated from the combination of the three short MCMC runs of this model, we retain the use of posterior probabilities derived from the three short runs for further discussions of the phylogeny. Estimates of model parameters derived from this long CNR-SSG MCMC run are given in Table 4.

Relative to the overall estimates of posterior probabilities (all 33 million generations minus 1 million burn-in), the deviation of posterior probability estimates at intervals of generations showed greater variance for the GTR+I+G model than did the CNR-SSG estimates (Fig. 9). The GTR+I+G model produced less precise point (intermediate interval) estimates than did the CNR-SSG model. In other words, as MCMC chains progressed through generations, the posterior probability estimates tended to vary more for the GTR+I+G than the CNR-SSG model. Although the GTR+I+G model included 20 nodes supported below 100%, whereas the CNR-SSG model included only 17, this bias was factored out by reporting deviations per interval after dividing by the number of nodes considered. This average nodal

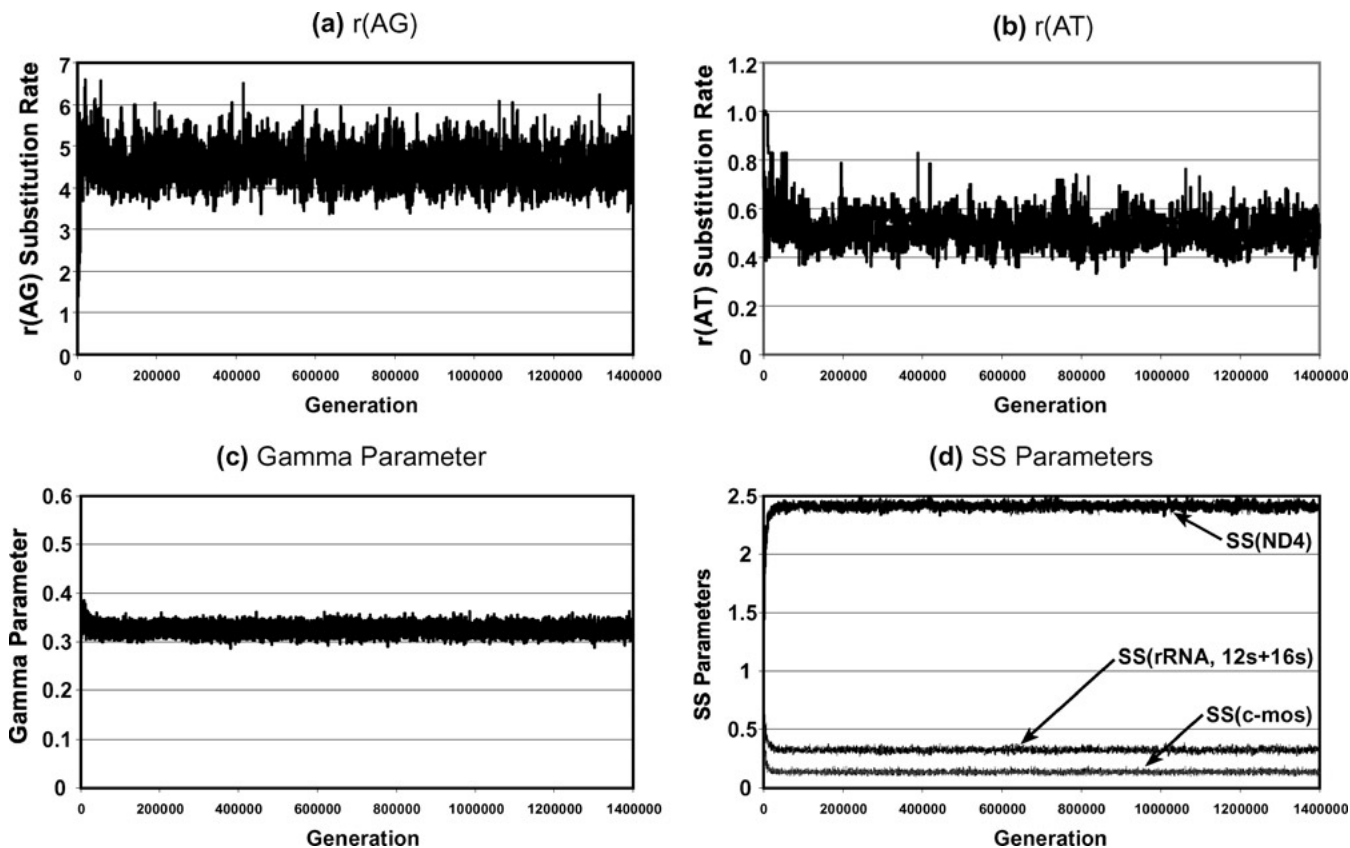


FIGURE 8. Plots of selected parameters of the CNR-SSG model through generations. All three independent runs are plotted per graph to show common burn-in rates and similar parameter estimates. (a) Plot of parametric estimates of  $r(A-G)$  from the GTR rate matrix. (b) Plot of the parametric estimates of  $r(A-T)$  from the GTR rate matrix. (c) Plot of the gamma parameter estimates. (d) Plot of the site-specific rate multiplier for the ND4, rRNA (12S+16S), and c-mos site specific partitions of the gamma parameter.

support deviation (from overall long run estimates) calculated from intervals of generations for the GTR+I+G model was more than twice that for the CNR-SSG model (Fig. 9). In comparisons of variance for nodes receiving similar levels of support (e.g., around 80% posterior probability), greater degrees of variation were evident in the GTR+I+G than the CNR-SSG model run (see Appendix 3 for detail, available at the Society of Systematic Biologists website, <http://systematicbiology.org>), suggesting that elevated deviation observed in GTR+I+G estimates were not particularly biased by overall higher posterior probability estimates from the CNR-SSG model. Despite variance in posterior probability estimates for intervals of generations, however, no latent trends were observed in posterior probabilities that may indicate that new tree islands were sampled only late in runs (after many generations) or that chains were not completely burned-in after the inferred burn-in period (based on likelihood plateau). Instead, fluctuations in nodal support through generations appear to represent oscillating patterns (see Appendix 3 for detail).

We reexamine the effect of model choice on the cumulative posterior probabilities for clades based on these

two extended MCMC runs (for the GTR+I+G and CNR-SSG models). Results from a Wilcoxon signed rank test returned very similar results as previous estimates based on the three short MCMC runs, suggesting the GTR+I+G model produced overall significantly lower estimates for posterior probabilities of clades than the preferred model (CNR-SSG;  $z = 2.334$ ,  $P = 0.019$ ).

The topology of the MCMC CNR-SSG tree (Fig. 6) has similarities with each of the other reconstructions but also differs from all aforementioned reconstructions in several ways. As with all of our phylogenetic reconstructions, *Alopoglossus* and *Ptychoglossus* form a well-supported clade sister to the rest of the Gymnophthalmidae. The Cercosaurinae is polyphyletic. As in the c-mos reconstruction, the Heterodactylini is paraphyletic with respect to the Gymnophthalmi. Although the placement of *Rhachisaurus* as the sister taxon to the Gymnophthalminae is unique among the parsimony and data partitions (mitochondrial and nuclear) in this study, it is in the same position as that was recovered by Pellegrino et al. (2001; their Fig. 4). Additionally, *Bachia* is recovered (with weak support) as the sister taxon to the Cercosaurini, as was found by Pellegrino et al. (2001).

TABLE 4. Parameter estimates for CNR-SSG model MCMC runs summarized as means with 95% credibility interval in parentheses.

	CNR-SSG run 1 (1.4 million generations)	CNR-SSG run 2 (1.4 million generations)	CNR-SSG run 3 (1.4 million generations)	CNR-SSG all short runs (1.4 million generations × 3)	CNR-SSG long run (33 million generations)
In likelihood	-30318.7 (-30338 to -30300.6)	-30317.6 (-30336.3 to -30302.2)	-30319.7 (-30340.0 to -30301.8)	-30318.7 (-30338.2 to -30301.5)	-30319.5 (-30335.6 to -30283.4)
pi(A)	0.393 (0.377-0.409)	0.393 (0.378-0.409)	0.393 (0.379-0.410)	0.393 (0.378-0.410)	0.393 (0.362-0.385)
pi(C)	0.283 (0.270-0.295)	0.283 (0.270-0.295)	0.283 (0.271-0.294)	0.283 (0.270-0.295)	0.283 (0.257-0.284)
pi(G)	0.087 (0.081-0.094)	0.087 (0.081-0.094)	0.087 (0.081-0.094)	0.087 (0.081-0.094)	0.087 (0.075-0.088)
pi(T)	0.237 (0.227-0.248)	0.237 (0.226-0.248)	0.237 (0.225-0.248)	0.237 (0.226-0.248)	0.237 (0.214-0.243)
r(A-C)	0.409 (0.312-0.491)	0.412 (0.339-0.506)	0.412 (0.325-0.509)	0.411 (0.324-0.504)	0.411 (0.269-0.400)
r(A-G)	4.491 (3.843-5.257)	4.483 (3.86-5.203)	4.489 (3.744-5.210)	4.488 (3.81-5.221)	4.502 (3.163-4.324)
r(A-T)	0.504 (0.403-0.614)	0.508 (0.414-0.630)	0.504 (0.391-0.622)	0.506 (0.404-0.624)	0.507 (0.333-0.490)
r(C-G)	0.364 (0.257-0.474)	0.358 (0.259-0.488)	0.357 (0.262-0.483)	0.359 (0.259-0.481)	0.360 (0.178-0.340)
r(C-T)	3.745 (3.139-4.423)	3.771 (3.199-4.469)	3.759 (3.120-4.503)	3.758 (3.152-4.466)	3.767 (2.651-3.509)
r(G-T)	1	1	1	1	1
Gamma parameter	0.326 (0.308-0.344)	0.3254 (0.308-0.343)	0.325 (0.308-0.343)	0.325 (0.308-0.343)	0.325 (0.291-0.329)
SS1 (c-mos)	0.134 (0.113-0.162)	0.134 (0.114-0.158)	0.133 (0.112-0.157)	0.134 (0.113-0.158)	0.133 (0.095-0.133)
SS2 (NID4)	2.409 (2.359-2.454)	2.414 (2.365-2.456)	2.415 (2.370-2.459)	2.412 (2.364-2.457)	2.412 (2.312-2.407)
SS3 (tRNA genes)	0.326 (0.295-0.362)	0.323 (0.293-0.356)	0.323 (0.290-0.354)	0.324 (0.292-0.358)	0.324 (0.262-0.329)

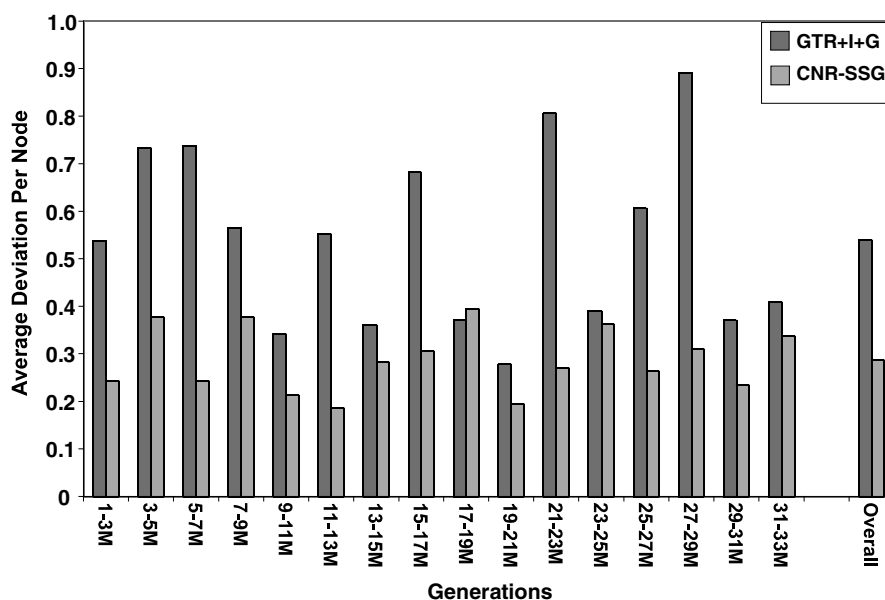


FIGURE 9. Comparison of deviation of posterior probability estimates at intervals of generations compared to overall means from long MCMC runs (33 million generations) for the GTR+I+G and CNR-SSG models. Values represent the absolute deviation of posterior probability estimates (relative to overall mean for long MCMC run) averaged across all nodes receiving less than 100% posterior probability.

#### Comparison Among Phylogenetic Reconstructions

Not all individual gene data sets (Fig. 2, and not shown) were in agreement with the combined tree (Fig. 6). The *c-mos* data partition agreed with the combined tree on higher-level relationships except for the placement of *Rhachisaurus* and *Bachia*. Similar to the parsimony reconstruction (Fig. 1), *Rhachisaurus* and *Bachia* formed a clade instead of *Rhachisaurus* being related to the Gymnophthalminae and *Bachia* to the Cercosaurini. The ND4 data partition supported a monophyletic Heterodactylini. The 16S data partition resolved the same general relationship as the combined data CNR-SSG tree except that one of the outgroups, *Tupinambis quadrilineatus*, was nested within the ingroup. The 12S data partition produced topologies most divergent from other genes, but nearly all of those relationships received poor posterior probability support.

## DISCUSSION

### Model Selection and Evaluation

Bayesian methods have greatly improved our ability to estimate phylogenies using larger datasets and complex models of evolution. However, this creates a seemingly paradoxical dilemma with regard to model complexity and overparameterization. In general, it is assumed that more realistic models of evolution will yield more accurate trees and clade credibility (posterior probability) values, thus perhaps favoring parameter-rich models, because interpretations of posterior probabilities are contingent on model specifications (Huelsenbeck et al., 2002). However, a key assumption of Wald's (1949) proof of the consistency of maximum likelihood

estimates is that all of the parameters of the likelihood function are identifiable from the true probability distribution of the data (Rogers, 2001). Even if a particular parameter may be intrinsic in the evolution of DNA sequences, we need to consider whether this parameter can be accurately estimated based on the data. This dilemma is manifested when attempting to construct and implement models that realistically describe DNA evolution, while avoiding overparameterization, or using more parameters than can be meaningfully estimated from the data.

In a Bayesian analysis, the problem of identifying the best model may be condensed to two intertwined issues: evaluating model performance and fit and examining the sensitivity of posterior probability distributions to model specifications (Gelman et al., 1995; Huelsenbeck et al., 2002). Detecting overparameterized models, however, is not readily accomplished, especially in a Bayesian phylogenetic framework (Huelsenbeck et al., 2002; Rannala, 2002). Several authors have suggested features of MCMC analyses that may be monitored to identify overparameterization, including poor convergence of MCMC chains (Carlin and Louis, 1996), a strong correlation among parameters in the posterior density despite independence under the prior density (Rannala, 2002), delayed convergence of a MCMC chain to a stationary plateau relative to less parameterized models (Rannala, 2002), and failure of multiple independent runs (chains) of the same model to converge on similar estimates of parameters and posterior probabilities (Huelsenbeck et al., 2002). We used these criteria, with the exception of testing among-parameter correlation, to guide the evaluation of what we tentatively

identified as the best-fit model (CNR-SSG). Testing for among-parameter correlation, in our case, was not possible because the nature of the model causes inherent correlation of the parameters of interest (those which partition the among-site rate variation across partitions).

Different evolutionary rates and among-site rate patterns may be intrinsic evolutionary characteristics of different genes owing to their genomic origin (organelle versus nuclear) or function (e.g., protein-coding versus non-protein-coding). Sufficient evidence exists to suggest that drastically different evolutionary rates and distinct, gene-specific among-site rates can be observed among different genes categorized within a particular class (e.g., among protein-coding mitochondrial genes; Miyata, 1982; Kelly and Rice, 1996). We used these observations to identify plausible alternative partitions across which to estimate specific rates of among-site rate variation. Comparison of post burn-in MCMC chain likelihood scores across alternative models for among-site rate variation showed that only the three partition (CNR-SSG and CNR-SAG) and four partition (4gene-SSG and 4gene-SAG) models fit the data better than the GTR+I+G model chosen by ModelTest. We found no evidence for substantial differences in burn-in time, chain likelihood score at stationarity, or overall clade posterior probability estimates across these models, yet we did detect a significant overall improvement in clade posterior probability estimates between one of these four models (CNR-SSG) and the GTR+I+G model. We found no evidence that this best-fit model (CNR-SSG) was parametrically over-fitted (excessively parameter rich). In fact, based on the analysis of the extended MCMC runs, we found this model to produce significantly more consistent posterior probabilities through generations than did the GTR+I+G model. Given available evidence, we concluded that the best-fit model of evolution, in keeping with our goal of practical improvement for the sake of phylogenetic inference, was the CNR-SSG model, upon which we base our preferred hypothesis for the phylogeny of Gymnophthalmidae.

Here, we summarize our approach to model construction and evaluation as an explicit hierarchical process:

1. Use hLRTs (e.g., ModelTest) to first identify best-fit conventional parameters (although other model choice criteria such as AIC [Akaike, 1974; also available in ModelTest], BIC [Schwarz, 1974], or DT [Minin et al., 2003] may be substituted)
2. Construct alternative models with data set partitions defined based on a priori expectations of potentially biologically relevant subsets of the data (e.g., protein-coding versus non-protein-coding genes or mitochondrial versus nuclear genes)
3. Examine model fit based on 95% CI of post burn-in MCMC chain likelihood values
4. Tentatively choose best-fit model
5. Examine this model for evidence of parameter identifiability or over-fitting
  - i. Compare relative burn-in period across alternative models
  - ii. Check for topological consistency across multiple runs of tentative model
  - iii. Examine consistency of parameter estimates across multiple independent runs of the tentatively optimal model
  - iv. Check for consistency of clade posterior probabilities across independent runs
  - v. Check for consistency of posterior probability estimates across generations for extended MCMC runs (with large number of generations)
6. Test for significant differences in posterior probabilities between tentative model and those models of similar fit to the data
7. Given evidence for parameter identifiability and significant changes to posterior probabilities, accept model. If identifiability is questionable or no significant changes to posterior probability are observed, reduce model parameterization and repeat model evaluation.

#### *Effects of Partitioning Gamma and Using Autocorrelated Rate Variation*

Several studies based on simulated data have strongly supported the view that maximum likelihood estimates of phylogeny remain accurate and robust even when the model used to estimate phylogeny differs markedly from that used to generate simulated data (Fukami-Kobayashi and Tatenno, 1991; Yang et al., 1994; Sullivan and Swofford, 2001). Our results support this conclusion using empirical data, in that several different models for among-site rate variation support the same or very similar topologies. Many authors have underscored the importance of including estimates of among-site rate variation (e.g., Yang, 1993; Sullivan and Swofford, 2001; Buckley and Cunningham, 2002; Nylander et al., 2004; see review in Yang, 1996b) in models of sequence evolution for increasing the consistency and accuracy of phylogenetic inference. Our results demonstrate that apparent inappropriate partitioning of gamma among loci (e.g., NM-SAG model) may lead to inconsistent and presumably inaccurate phylogenetic inferences. The fact that our preferred model (CNR-SSG) provided a significant increase in overall posterior probability estimates for clades over the GTR+I+G model suggests that well fitted partitioning of among-site rate variation appears to significantly affect the posterior probability distributions of MCMC analyses. These results parallel previous studies that have demonstrated the significant effects of substitution model on maximum likelihood bootstrap support (Yang et al., 1995; Sullivan et al., 1997; Buckley et al., 2001; Buckley and Cunningham, 2002).

An interesting, yet difficult to interpret, result was observed in comparisons of posterior probability monitored through intervals during extended MCMC runs. We found the CNR-SSG model to produce more consistent posterior probabilities through generations than did



the GTR+I+G model. If we assume that the complexity of tree space remained relatively constant under the two models, this may suggest that the MCMC chains of the CNR-SSG model were more consistent over intervals of generations with respect to the regular visitation of tree islands. Alternatively, it seems possible that implementing the partitioned gamma model (CNR-SSG) may have reduced the complexity of tree space by decreasing the number of optimal or near optimal peaks (reducing the number of major islands visited by MCMC chains over time), thereby reducing the variance through generations in trees sampled in the posterior distribution. This may have been accomplished by reducing the likelihood of certain peaks within tree space due to the different parameterization of among-site rates, thereby decreasing the number of near optimal tree islands. In general, the properties associated with the behavior of MCMC chains in tree space through generations has been essentially untouched in the literature, yet represents a significant gap in our understanding of Bayesian MCMC analyses. Future research is clearly necessary to answer questions about the number of generations and independent runs required for robust conclusions from MCMC and also how this may relate to model complexity and changes to the general topology of tree space.

Although the results of this study favor use of models that partition among-site rate variation, they also highlight a potential pitfall of such parameter-rich models. Not all alternative models improved model fit relative to GTR+I+G. The GTR+AG, PR-SSG, PR-SAG, NM-SSG, and NM-SAG models all decreased the fit of the model to the data, relative to the GTR+I+G model (chosen initially by hLRT criteria). These results re-emphasize the need to test the fit of alternative models instead of choosing a particular model a priori (e.g., Huelsenbeck and Crandall, 1997; Posada and Crandall, 2001; Minin et al., 2003). The GTR+I+G model not only fit the data better than some partitioned models (e.g., NM-SSG, PR-SAG; Fig. 4), but also recovered the identical topology while (on average) underestimating posterior probability support for clades (Fig. 7). These findings support the utility of this conventionally employed model and suggest that previous analyses using this model are likely to be as robust (with regard to topology) as more complex models, but provide more conservative estimates of posterior probability support for clades. However, our analysis of extended MCMC runs suggests that, for a reason that is not immediately clear, the GTR+I+G model appeared to take a large number of generations to undergo oscillation cycles with respect to estimates of posterior probabilities. This suggests that, for some models, at least one extended MCMC run (with a large number of generations) is desirable to precisely and accurately estimate posterior probabilities so that trees are sampled in the posterior distribution according to their posterior probability (Swofford, Warren, and Wilgenbusch, unpublished data). It is encouraging, from the standpoint of computational feasibility, that the estimates of model parameters, chain likelihood scores, and, particularly, posterior probabilities derived from the combination of three short

(1.4 million generation) independent MCMC runs provided what appears to be, at least, a sufficient approximation of posterior probabilities derived from much longer MCMC runs.

Several authors have demonstrated the utility of employing a parameter to account for autocorrelated among-site rate variation in phylogenetic analyses (e.g., Yang, 1995; Penny et al., 2001; Huelsenbeck, 2002). Although evidence for the occurrence of autocorrelated rates has been well documented (Yang, 1995; Nielsen, 1997; Penny et al., 2001), we found the addition of this parameter to alternative models to be of limited value for improving the fit of models to our data. As can be seen in Figure 4, models which fit the data poorly (PR-SSG and NM-SSG) did appear to be notably improved with the addition of a parameter for autocorrelation of among-site rates, although this increase in fit did not exceed that of the GTR+I+G model (or the CNR or 4gene models). Among models which showed the best-fit to the data (CNR and 4gene), the addition of a parameter to account for among-site rate autocorrelation only slightly increased the likelihood scores for MCMC chains such that there was still broad overlap in the 95% CI of likelihood scores (Fig. 4). Similarly, Wilcoxon signed rank tests comparing overall posterior probabilities for clades between SSG and SAG variants of the CNR and 4gene models found no significant differences attributable to the addition of the auto-correlation parameter.

In this study we have concentrated on accounting for one particular type of heterogeneity in among-site rate patterns in combined DNA sequence analysis, that which exists at or above the level of a gene or locus, ignoring potential partitions that may be prescribed within genes. Models that do examine and attempt to account for within gene heterogeneity by constructing partitions based on codon position (e.g., Yang, 1996a; Krajewski et al., 1999; Buckley et al., 2001), protein domain (Herron et al., in press), or secondary structure for rRNA or tRNA genes (e.g., Schoniger and von Haeseler, 1994; Savill et al., 2001) have also been implemented. These intralocus partitions have yet to be thoroughly evaluated in a Bayesian framework and may potentially add additional realistic parameters to models of sequence evolution, especially in cases where very distant relationships are inferred (Penny et al., 2001) or where extreme accuracy of branch length estimates or model parameters are particularly critical to conclusions (Yang et al., 1994). Understanding and testing of parametric identifiability in complex models have been poorly studied and clearly requires additional attention. This issue, in addition to topology and posterior probability sensitivity to model choice, would benefit from future investigations using both simulated data and known phylogenies where more definitive conclusions about the effects of model choice may be drawn.

#### *Taxonomic Considerations and Alterations*

Much of our phylogeny reconstruction is consistent with that recovered by Pellegrino et al. (2001). However,

our preferred phylogenetic hypothesis (combined data MCMC CNR-SSG reconstruction; Fig. 6) suggests four higher level taxonomic changes to the current classification (Pellegrino et al., 2001). The first change is that *Ptychoglossus* appears to be most closely related to *Alopoglossus* and not to the Cercosaurini. The placement of *Ptychoglossus* in the Cercosaurini by Pellegrino et al. (2001) was presumably the result of the swapping of taxon names between *Ptychoglossus* and *Neusticurus juruazensis*, as discussed in Appendix 2. This relationship was also inferred from the nuclear partition trees of Pellegrino et al. (2001) (and the c-mos reconstruction of Harris, 2003) in which *Ptychoglossus* was sister to the three *Alopoglossus* species. After making the correction to the Pellegrino et al. (2001) data set, and adding our own sequences for this taxon, it seems clear that *Ptychoglossus brevifrontalis* is sister to *Alopoglossus*; therefore, we remove *Ptychoglossus* from the Cercosaurinae and place it in the Alopoglossinae. This relationship is also supported by the morphological synapomorphy (present in both *Ptychoglossus* and *Alopoglossus*) of infralingual plicae, unique in the family Gymnophthalmidae.

The second taxonomic alteration involves the tribe Heterodactylini. This tribe is paraphyletic with respect to the Gymnophthalmini in our combined tree (Fig. 6) and in the c-mos (Fig. 2) and 16S (not shown) reconstructions. The paraphyly of the tribe was also apparent in Pellegrino et al.'s (2001) maximum likelihood tree. Because there does not appear to be sufficient support for recognizing a separate tribe Heterodactylini, we remove both of the Gymnophthalmini and Heterodactylini tribal names and refer all of the pertaining genera to subfamily Gymnophthalminae with no tribes.

The third taxonomic alteration involves species belonging to the cercosaurine tribe Ecleopini. The CNR-SSG tree (Fig. 6) suggests that the ecleopiines and the cercosauriines do not comprise a monophyletic Cercosaurinae. Although posterior probability support for intervening clades is low, monophyly of both groups is well supported. The Ecleopini appears to be distantly related to the Cercosaurini and we hereby raise the status of the former members of tribe Ecleopini (*Amapasaurus*, *Anotosaura*, *Arthrosaura*, *Colobosauroides*, *Ecleopus*, and *Leposoma*; Pellegrino et al., 2001) to subfamily status, the Ecleopinae Fitzinger.

The fourth taxonomic alteration involves the placement of *Bachia*. Pellegrino et al. (2001) recovered its placement as basal within the Cercosaurini. The node joining *Bachia* to the rest of the Cercosaurini was supported by bootstrap values less than 50% on their parsimony tree and by 81% on their maximum likelihood tree. We found conflict between our parsimony reconstructions and Bayesian reconstructions. In the parsimony trees (and 16S and c-mos individual Bayesian gene trees; 16S not shown) we found *Bachia* to be closely related to *Rhachisaurus brachylepis*, either joined with the Ecleopini or distantly related to the Cercosaurinae. In our CNR-SSG reconstruction *Bachia* appears to be the sister lineage to the rest of the Cercosaurini with low posterior probabilities supporting *Bachia* in that position. In addition,

a large genetic distance separated *Bachia* from the other cercosauriines. Based on these data we are still unsure of the phylogenetic placement of *Bachia* within the family. However, we are confident that *Bachia* appears to be distantly related to all other sampled taxa. We believe the best course of action at the present time is to leave *Bachia* in the Cercosaurinae but elevate the genus to tribe status, the Bachini. In this way the relationships of this genus with other genera of the Cercosaurinae are not confused.

A new phylogenetic classification for the family is presented in Table 5. Based on the additions we made to the Pellegrino et al. (2001) data set, several other taxonomic issues warrant mention. The addition of *Pholidobolus*

TABLE 5. Current phylogenetic classification of family Gymnophthalmidae.

Taxon
Gymnophthalmidae Merrem, 1820
Alopoglossinae Pellegrino, Rodrigues, Yonenaga-Yassuda, and Sites, 2001
<i>Alopoglossus</i> Boulenger, 1885
<i>Ptychoglossus</i> Boulenger, 1890
Cercosaurinae Gray, 1838
Tribe Bachini New Tribe
<i>Bachia</i> Gray, 1845
Tribe Cercosaurini Gray, 1838
<i>Anadia</i> Gray, 1845
<i>Cercosaura</i> Wagler, 1830
<i>Echinosaura</i> Boulenger, 1890
<i>Euspondylus</i> Tschudi, 1845
<i>Macropholidus</i> Noble, 1921
<i>Neusticurus</i> Duméril and Bibron, 1839
<i>Opiputeer</i> Uzzell, 1969
<i>Pholidobolus</i> Peters, 1862
<i>Placosoma</i> Tschudi, 1847
<i>Proctoporus</i> Tschudi, 1845
<i>Riolama</i> Uzzell, 1973
<i>Teuchocercus</i> Fritts and Smith, 1969
Ecleopinae Fitzinger, 1843
<i>Amapasaurus</i> Cunha, 1970
<i>Anotosaura</i> Amaral, 1933
<i>Arthrosaura</i> Boulenger, 1885
<i>Colobosauroides</i> Cunha and Lima Verde, 1991
<i>Ecleopus</i> Duméril and Bibron, 1839
<i>Leposoma</i> Spix, 1825
Gymnophthalminae Merrem, 1820
<i>Calyptommatius</i> Rodrigues, 1991
<i>Colobodactylus</i> Amaral, 1933
<i>Colobosaura</i> Boulenger, 1887
<i>Heterodactylus</i> Spix, 1825
<i>Iphisa</i> Gray, 1851
<i>Gymnophthalmus</i> Merrem, 1820
<i>Micrablepharus</i> Dunn, 1932
<i>Nothobachia</i> Rodrigues, 1984
<i>Procellosaurinus</i> Rodrigues, 1991
<i>Psilophthalmus</i> Rodrigues, 1991
<i>Stenolepis</i> Boulenger, 1888
<i>Tretioscincus</i> Cope, 1862
<i>Vanzosaura</i> Rodrigues, 1991
Rhachisaurinae Pellegrino, Rodrigues, Yonenaga-Yassuda, and Sites, 2001
<i>Rhachisaurus</i> Pellegrino, Rodrigues, Yonenaga-Yassuda, and Sites, 2001

Note. We are unable to place *Adercosaurus* Myers and Donnelly, 2001, definitively in a subfamily because we were unable to examine specimens of that genus. Based on the information provided by Myers and Donnelly (2001), this genus may belong in the Alopoglossinae, the Ecleopinae, or the Cercosaurini.

*macbrydei* provides preliminary support for a monophyletic *Pholidobolus*. The newly redesignated genus *Cercosaura*, which now includes all taxa formerly placed in *Pantodactylus* and *Prionodactylus* (Doan, 2003a), is supported in this study. The addition of *Neusticurus strangulatus* shows that this species forms a clade with two other members of its genus, *N. ecleopus* and *N. juruazensis*, but overall the genus is polyphyletic. Additionally, *Anotosaura* is paraphyletic with respect to *Colobosauroides* and *Colobosaura* is paraphyletic with respect to *Iphisa*.

*Proctoporus* is the genus that was not included by Pellegrino et al. (2001). Contrary to the conclusions made by Doan (2003b) using morphological data, *Proctoporus* appears to be a polyphyletic member of the Cercosaurini. In the CNR-SSG reconstruction two separate *Proctoporus* clades are apparent, separated from each other by *Pholidobolus*, *Cercosaura*, and one clade of *Neusticurus*. One *Proctoporus* clade is composed of members from Ecuador, whereas the other includes members from Peru and Bolivia. In the parsimony reconstruction *Proctoporus ventrimaculatus* additionally forms a third lineage that appears to be most closely related to *Cercosaura*. This species (including an unidentified specimen designated as *P. cf. ventrimaculatus*) is the sole species from northern Peru, separated by a vast distance from the Ecuadorian clade to its north and the southern Peruvian and Bolivian clade to its south. It is clear that taxonomic rearrangement is necessary to rectify the taxonomy of this genus. This work is underway by Doan and Castoe (in preparation).

#### ACKNOWLEDGMENTS

We would like to thank L. Ford and C. Raxworthy of AMNH and L. Trueb of KU for the donation of tissues. We thank D. Calloway, P. Chippindale, M. Herron, B. Noonan, C. O'Reilly, A. Rambout, R. Ruggiero, C. Simon, J. Sites, E. Smith, J. Sullivan and two anonymous reviewers for helpful methodological suggestions and comments on the manuscript. We thank the Biology Department at the University of Central Florida, J. Fauth, and J. Weishampel for providing computer access.

#### REFERENCES

- Akaike, H. 1974. A new look at statistical model identification. *IEE Trans. Autom. Contr.* 19:716–723.
- 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.
- Arévalo, E. S., S. K. Davis, and J. W. Sites Jr. 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.
- Buckley, T. R. 2002. Model misspecification and probabilistic tests of topology: Evidence from empirical data sets. *Syst. Biol.* 51:509–523.
- 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.
- Buckley, T. R., C. Simon, and G. K. Chambers. 2001. Exploring among-site rate variation models in a maximum likelihood framework using empirical data: Effects of model assumptions on estimates of topology, branch lengths, and bootstrap support. *Syst. Biol.* 50:67–86.
- Carlin, B. P., and T. A. Louis. 1996. Bayes and empirical Bayes methods for data analysis. Chapman and Hall, New York.
- Caterino, M. S., R. D. Reed, M. M. Kuo, and F. A. H. Sperling. 2001. A partitioned likelihood analysis of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). *Syst. Biol.* 50:106–127.
- Cummings, M. P., S. A. Handley, D. S. Meyers, D. L. Reed, A. Rokas, and K. Winka. 2003. Comparing bootstrap and posterior probability values in the four-taxon case. *Syst. Biol.* 52:477–487.
- Dixon, M. T., and D. M. Hillis. 1993. Ribosomal RNA secondary structure: Compensatory mutations and implications for phylogenetic analysis. *Mol. Biol. Evol.* 10:256–267.
- Doan, T. M. 2003a. A new phylogenetic classification for the gymnophthalmid genera *Cercosaura*, *Pantodactylus*, and *Prionodactylus* (Reptilia: Squamata). *Zool. J. Linn. Soc.* 137:101–115.
- Doan, T. M. 2003b. A south-to-north biogeographic hypothesis for Andean speciation: Evidence from the lizard genus *Proctoporus* (Reptilia, Gymnophthalmidae). *J. Biogeography* 30:361–374.
- Doan, T. M., and T. A. Castoe. 2003. Using morphological and molecular evidence to infer species boundaries within *Proctoporus bolivianus* Werner (Squamata: Gymnophthalmidae). *Herpetologica* 59:433–450.
- Douady, C. J., F. Delsuc, Y. Boucher, W. F. Doolittle, and E. J. P. Douzery. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol. Biol. Evol.* 20:248–254.
- Erixon, S. P., B. Britton, and B. Oxelman. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Syst. Biol.* 52:665–673.
- Felsenstein, J. 1968. Statistical inference and the estimation of phylogenies. University of Chicago, Chicago.
- Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- Fukami-Kobayashi, K., and Y. Tateno. 1991. Robustness of maximum likelihood tree estimation against different patterns of base substitutions. *J. Mol. Evol.* 32:79–91.
- Gasser, R. B., X. Zhu, N. B. Chilton, L. A. Newton, T. Nedergaard, and P. Guldberg. 1998. Analysis of sequence homogenization in rDNA arrays of *Haemonchus contortus* by denaturing gradient gel electrophoresis. *Electrophoresis* 19:2391–2395.
- Gene Codes Corp. 1996. Sequencher, version 3.1. Gene Codes, Ann Arbor.
- Gonzales, I. L., and J. E. Sylvester. 2001. Human rRNA: Evolutionary patterns within genes and tandem arrays derived from multiple chromosomes. *Genomics* 73:255–263.
- Gutell, R. R. 1994. Collection of small subunit (16S and 16S-like) ribosomal RNA structures. *Nucleic Acid Res.* 22:3502–3507.
- Gutell, R. R., N. Larsen, and C. R. Woese. 1994. Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol. Rev.* 58:10–26.
- Harris, D. J. 2003. Codon bias variation in *C-mos* between squamate families might distort phylogenetic inferences. *Mol. Phylogenet. Evol.* 27:540–544.
- Herron, M. D., T. A. Castoe, and C. L. Parkinson. in press. Sciurid phylogeny and the paraphyly of Holarctic ground squirrels (*Spermophilus*). *Mol. Phylogenet. Evol.*
- Hickson, R., C. Simon, A. Cooper, G. Spicer, J. Sullivan, and D. Penny. 1996. Conserved sequence motifs, alignment, and secondary structure for the third domain of animal 12s rRNA. *Mol. Biol. Evol.* 13:150–169.
- Hillis, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Syst. Biol.* 47:3–8.
- Hillis, D. M., and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42:182–192.
- Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. *Syst. Biol.* 44:17–48.
- Huelsenbeck, J. P. 1997. Is the Felsenstein zone a fly trap? *Syst. Biol.* 46:69–74.
- Huelsenbeck, J. P. 2002. Testing a covarion model of DNA substitution. *Mol. Biol. Evol.* 19:698–707.
- Huelsenbeck, J. P., and K. A. Crandall. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Ann. Rev. Ecol. Syst.* 28:437–466.
- Huelsenbeck, J. P., B. Larget, R. Miller, and F. Ronquist. 2002. Potential applications and pitfalls of bayesian inference of phylogeny. *Syst. Biol.* 51:673–688.

- Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17:754–755.
- Huelsenbeck, J. P., R. Ronquist, R. Nielsen, and J. Bollback. 2001. Bayesian inference of phylogeny and its impacts on evolutionary biology. *Science* 294:2310–2314.
- International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860–921.
- Kelly, C., and J. Rice. 1996. Modeling nucleotide evolution: A heterogeneous rate analysis. *Math. Biosci.* 133:85–109.
- Kimura, M. 1986. The role of compensatory neutral mutations in molecular evolution. *J. Genetics* 64:7–19.
- Krajewski, C., M. G. Fain, L. Buckley, and D. G. King. 1999. Dynamically heterogeneous partitions and phylogenetic inference: An evaluation of analytical strategies with cytochrome b and ND6 gene sequences in cranes. *Mol. Phylogenet. Evol.* 13:302–313.
- 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.
- Long, E. O., and I. B. Dawid. 1980. Repeated genes in eukaryotes. *Ann. Rev. Biochem.* 49:727–764.
- Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogeny estimation. *Syst. Biol.* 52:674–683.
- Miyata, T. 1982. Evolutionary changes and functional constraints in DNA sequences. Pages 233–266 in *Molecular evolution, protein polymorphism and the neutral allele theory* (M. Kimura, ed.). Japan Scientific Press, Tokyo.
- Moncalvo, J. M., D. Drehtmel, and R. Vilgalys. 2000. Variation in modes and rates of evolution in nuclear and mitochondrial ribosomal DNA in the mushroom genus *Aminita* (Agaricales, Basidiomycota): Phylogenetic implications. *Mol. Phylogenet. Evol.* 16:48–63.
- Muse, S. V. 1995. Evolutionary analyses when nucleotides do not evolve independently. Pages 115–124 in *Current topics on molecular evolution* (M. Nei and N. Takahata, eds.). Institute on Molecular Evolution and Genetics, University Park, PA.
- Myers, C. W., and M. A. Donnelly. 2001. Herpetofauna of the Yutajé-Corocoro Massif, Venezuela: Second report from the Robert G. Goelet American Museum-Terramar expedition to the northwestern tepuis. *Bull. Am. Mus. Nat. Hist.* 261:1–85.
- Nielsen, R. 1997. Site-by-site estimation of the rate of substitution and the correlation of rates in mitochondrial DNA. *Syst. Biol.* 46:346–353.
- Nicholas, K. B., and H. B. Nicholas, Jr. 1997. GeneDoc: A tool for editing and annotating multiple sequence alignments. Distributed by the author at: [www.cris.com/~Ketchup/genedoc.shtml](http://www.cris.com/~Ketchup/genedoc.shtml).
- Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* 53:47–67.
- Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* 5:568–583.
- Pardue, M. L. 1974. Localization of repeated DNA sequences in *Xenopus* chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* 38:475–482.
- Parkinson, C. L. 1999. Molecular systematics and biogeographical history of pitvipers as determined by mitochondrial ribosomal DNA sequences. *Copeia* 1999:576–586.
- Parkinson, C. L., S. M. Moody, and J. E. Alquist. 1997. Phylogenetic relationships of the 'Agkistrodon Complex' based on mitochondrial DNA sequence data. Pages 63–78 in *Venomous snakes: Ecology, evolution, and snakebite* (R. S. Thorpe, W. Wüster, and A. Malhotra, eds.). Symposia of the Zoological Society of London. Clarendon Press, Oxford.
- Pellegrino, K. C. M., M. T. Rodrigues, Y. Yonenaga-Yassuda, and J. W. Sites Jr. 2001. A molecular perspective on the evolution of microteiid lizards (Squamata: Gymnophthalmidae), and a new classification for the family. *Biol. J. Linn. Soc.* 74:315–338.
- Penny, D., B. J. McComish, M. A. Carleston, and M. D. Hendy. 2001. Mathematical elegance with biochemical realism: The covarion model of molecular evolution. *J. Mol. Evol.* 53:711–723.
- Posada, D., and K. A. Crandall. 1998. Model-Test: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Posada, D., and K. A. Crandall. 2001. Selecting the best-fit model of nucleotide substitution. *Syst. Biol.* 50:580–601.
- Pupko, T., D. Huchon, Y. Cao, N. Okada, and M. Hasegawa. 2002. Combining multiple data sets in a likelihood analysis: Which models are the best? *Mol. Biol. Evol.* 19:2294–2307.
- Rannala, B. 2002. Identifiability of parameters in MCMC Bayesian inference of phylogeny. *Syst. Biol.* 51:754–760.
- Rogers, J. S. 2001. Maximum likelihood estimation of phylogenetic trees in consistent when substitution rates vary according to the invariable sites plus gamma distribution. *Syst. Biol.* 50:713–722.
- Saint K. M., C. C. Austin, S. C. Donnellan, and M. N. Hutchinson. 1998. C-mos, a nuclear marker useful for squamate phylogenetic analysis. *Mol. Phylogenet. Evol.* 10:259–263.
- Savill, N. J., D. C. Hoyle, and P. G. Higgs. 2001. RNA sequence evolution with secondary structure constraints: Comparison of substitution rate models using maximum-likelihood methods. *Genetics* 157:399–411.
- Schöniger, M., and A. von Haeseler. 1994. A stochastic model and the evolution of autocorrelated DNA sequences. *Mol. Phylogenet. Evol.* 3:240–247.
- Schwarz, G. 1974. Estimating the dimension of a model. *Ann. Stat.* 6:461–464.
- Simon, C., F. Frati, A. Beckenbach, B. Brespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved PCR primers. *Ann. Entomol. Soc. Am.* 87:651–701.
- StatSoft, Inc. 1993. Statistica for Windows, release 4.5. StatSoft, Tulsa.
- Sullivan, J., J. A. Markert, and C. William Kilpatrick. 1997. Phylogeography and molecular systematics of the *Peromyscus aztecus* species group (Rodentia: Muridae) inferred using parsimony and likelihood. *Syst. Biol.* 46:426–440.
- Sullivan, J., and D. L. Swofford. 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? *Syst. Biol.* 50:723–729.
- Suzuki, Y., G. V. Glazko, and M. Nei. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proc. Natl. Acad. Sci. U.S.A.* 99:16138–16143.
- Swofford, D. L. 2002. PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods), version 4.0b10, Sinauer Associates, Sunderland, MA.
- Tavaré, S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Pages 57–86 in *Some mathematical questions in biology—DNA sequence analysis* (R. M. Miura, ed.). American Math Society, Providence, RI.
- Titus, T. A., and D. R. Frost. 1996. Molecular homology assessment and phylogeny in the lizard family Opluridae (Squamata: Iguania). *Mol. Phylogenet. Evol.* 6:49–62.
- Wald, A. 1949. Note on the consistency of the maximum likelihood estimate. *Ann. Math. Stat.* 20:595–601.
- Wilcox, T. P., D. J. Zwickl, T. A. Heath, and D. M. Hillis. 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. *Mol. Phylogenet. Evol.* 25:361–371.
- Wu, C. I. 1991. Inferences of species phylogeny in relation to segregation of ancient polymorphisms. *Genetics* 127:429–435.
- Yang, Z. 1993. Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites: Approximate methods. *Mol. Biol. Evol.* 10:1396–1401.
- Yang, Z. 1995. A space-time process model for the evolution of DNA sequences. *Genetics* 139:993–1005.
- Yang, Z. 1996a. Maximum likelihood models for combined analyses of multiple sequence data. *J. Mol. Evol.* 42:587–596.
- Yang, Z. 1996b. Among-site rate variation and its impacts on phylogenetic analyses. *Trends Ecol. Evol.* 11:367–372.
- Yang, Z., N. Goldman, and A. Friday. 1994. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. *Mol. Biol. Evol.* 11:316–324.
- Yang, Z., N. Goldman, and A. Friday. 1995. Maximum likelihood trees from DNA sequences: A peculiar statistical estimation problem. *Syst. Biol.* 44:384–399.

First submitted 18 May 2003; reviews returned 15 March 2003;

final acceptance 14 December 2003

Associate Editor: Jack Sites

## APPENDIX 1

List of operational taxonomic units (OTUs) used in this study with GenBank accession numbers. Cells with an X indicate that gene sequence was not used in this study. (a) and (b) refer to individuals indicated in the figures. Museum accession numbers for specimens sequenced in this study are given. Acronyms for museums are KU (University of Kansas), MHNSM (Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Lima, Peru), QCAZ (Museo de Zoología, Pontificia Universidad Católica del Ecuador, Quito, Ecuador), and UTA (University of Texas at Arlington).

OTU	Museum number	ND4	c-mos	12S	16S
<i>Alopoglossus atriventris</i>		AF420908	AF420821	AF420695	AF420746
<i>Alopoglossus carinicaudatus</i>		AF420909	AF420847	AF420693	AF420744
<i>Alopoglossus copii</i>		AF420865	AF420819	AF420692	AF420745
<i>Anotosaura spn</i>		AF420902	X	AF420682	AF420719
<i>Anotosaura vanzolinia</i>		AF420910	X	AF420670	AF420724
<i>Arthrosaura kockii</i>		AF420866	X	AF420680	AF420721
<i>Arthrosaura reticulata</i>		AF420894	X	AF420676	AF420722
<i>Bachia bresslaui</i>		AF420876	AF420860	X	AF420755
<i>Bachia dorbignyi</i>		AF420892	X	AF420688	AF420754
<i>Bachia flavescens</i>		AF420869	AF420859	AF420705	AF420753
<i>Calyptommatius leiolepis</i>		AF420874	AF420858	AF420683	AF420712
<i>Calyptommatius nicterus</i>		AF420903	AF420822	AF420684	AF420747
<i>Calyptommatius sinebrachiatus</i>		AF420873	AF420832	AF420685	AF420720
<i>Cercosaura argulus</i> (a)		AF420896	AF420838	AF420698	AF420751
<i>Cercosaura argulus</i> (b)		AF420893	AF420852	AF420696	AF420750
<i>Cercosaura eigenmanni</i>		AF420895	AF420828	AF420690	AF420728
<i>Cercosaura ocellata</i>		AF420883	AF420834	AF420677	AF420731
<i>Cercosaura quadrilineata</i>		AF420880	AF420830	AF420672	AF420717
<i>Cercosaura schreibersii albostrigata</i>		AF420882	AF420856	AF420658	AF420729
<i>Cercosaura schreibersii schreibersii</i>		AF420911	AF420817	AF420686	AF420749
<i>Colobodactylus dalcyanus</i>		AF420881	AF420844	AF420663	AF420736
<i>Colobodactylus taunayi</i>		X	AF420831	AF420662	AF420741
<i>Colobosaura mentalis</i>		AF420899	AF420842	AF420694	AF420726
<i>Colobosaura modesta</i>		AF420887	AF420845	AF420666	AF420733
<i>Colobosaura spn</i>		AF420868	AF420840	AF420667	AF420739
<i>Colobosauroides cearensis</i>		AF420886	AF420849	AF420659	AF420727
<i>Ecpleopus gaudichaudii</i>		AF420901	AF420855	AF420660	AF420738
<i>Gymnophthalmus leucomystax</i>		AF420906	AF420824	AF420675	AF420715
<i>Gymnophthalmus vanzoi</i>		AF420867	AF420827	AF420687	AF420743
<i>Heterodactylus imbricatus</i>		AF420885	AF420835	AF420661	AF420725
<i>Iphisa elegans</i>		AF420889	AF420843	AF420668	AF420714
<i>Leposoma osvaldoi</i>		AF420897	AF420854	AF420678	AF420723
<i>Leposoma percarinatum</i>		AF420898	X	AF420700	AF420735
<i>Micrablepharus atticolus</i>		AF420904	AF420826	AF420664	AF420718
<i>Micrablepharus maximiliani</i>		AF420875	AF420850	AF420657	AF420730
<i>Neusticurus bicarinatus</i>		X	AF420816	AF420671	AF420708
<i>Neusticurus ecpleopus</i>		AF420890	AF420829	AF420656	AF420748
<i>Neusticurus juruaensis</i>		AF420878	AF420857	AF420704	AF420758
<i>Neusticurus rudis</i>		AF420905	X	AF420689	AF420709
<i>Neusticurus strangulatus</i>	KU 21677	AY507885	X	AY507847	AY507866
<i>Nothobachia ablephara</i>		AF420900	AF420851	AF420669	AF420740
<i>Pholidobolus macbrydei</i>	KU 218406	AY507886	AY507896	AY507848	AY507867
<i>Pholidobolus montium</i>		AF420884	AF420820	AF420701	AF420756
<i>Placosoma cordylinum</i>		AF420879	AF420823	AF420673	AF420734
<i>Placosoma glabellum</i>		AF420907	AF420833	AF420674	AF420742
<i>Procellosaurinus erythrocerus</i>		AF420870	AF420836	AF420679	AF420711
<i>Procellosaurinus tetradactylus</i>		AF420871	AF420818	AF420703	AF420713
<i>Proctoporus bolivianus</i> (a)	UTA R-51506	AY225175	AY507898	AY507851	AY507869
<i>Proctoporus bolivianus</i> (b)	UTA R-51487	AY225180	AY507897	AY507850	AY507868
<i>Proctoporus cashcaensis</i>	KU 217205	AY507887	X	AY507852	AY507870
<i>Proctoporus colomaromani</i>	KU 217209	AY507888	AY507899	AY507853	AY507871
<i>Proctoporus guentheri</i> (a)	UTA R-51515	AY225185	AY507900	AY507849	AY507872
<i>Proctoporus guentheri</i> (b)	UTA R-51517	AY225169	AY507901	AY507854	AY507873
<i>Proctoporus orcesi</i>	KU 221772	AY507889	X	AY507855	AY507874
<i>Proctoporus simoterus</i>	KU 217207	AY507890	AY507902	AY507856	AY507875
<i>Proctoporus succullucu</i> (a)	UTA R-51478	AY225171	AY507903	AY507857	AY507878
<i>Proctoporus succullucu</i> (b)	UTA R-51496	AY225177	AY507904	AY507858	AY507879
<i>Proctoporus unicolor</i>	KU 217211	AY507893	AY507907	AY507862	AY507880
<i>Proctoporus unsaaciae</i> (a)	UTA R-51477	AY225170	AY507909	AY507860	AY507881
<i>Proctoporus unsaaciae</i> (b)	UTA R-51488	AY225186	AY507908	AY507859	AY507882
<i>Proctoporus ventrimaculatus</i>	KU 219838	AY507894	AY507910	AY507863	AY507883
<i>Proctoporus cf. ventrimaculatus</i>	KU 212687	AY507891	AY507903	AY507864	AY507876
<i>Proctoporus sp. K19</i>	QCAZ 879	AY507892	AY507904	AY507861	AY507877

(Continued on next page)

## APPENDIX 1 (CONTINUED)

OTU	Museum number	ND4	c-mos	12S	16S
<i>Psilophthalmus paeminosus</i>		AF420872	AF420825	AF420702	AF420710
<i>Ptychoglossus brevifrontalis</i>		X	AF420848	AF420697	AF420757
<i>Ptychoglossus brevifrontalis</i>	MHNSM	AY507895	AY507911	AY507865	AY507884
<i>Rhachisaurus brachylepis</i>		AF420877	AF420853	AF420665	AF420737
<i>Tretioscincus agilis</i>		AF420891	AF420837	AF420681	AF420732
<i>Tretioscincus oriximinensis</i>		AF420888	AF420846	AF420691	AF420752
<i>Vanzosaura rubricauda</i>		X	AF420839	AF420699	AF420716
<i>Cnemidophorus ocellifer</i>		AF420914	AF420862	AF420706	AF420759
<i>Kentropyx calcarata</i>		AF420913	AF420864	AF420707	AF420760
<i>Tupinambis quadrilineatus</i>		AF420912	AF420863	X	AF420761

## APPENDIX 2

PROBLEMS DISCOVERED WITH THE PELLEGRINO ET AL.  
(2001) DATA SET

After reconstructing phylogenies of each data partition separately and combined, we identified four taxa from the Pellegrino et al. (2001) study that demonstrated remarkable incongruence with regard to topological placement among individual gene phylogenies. After careful examination of sequences, phylogenetic trees for independent data partitions, and verification based on our own data collection, we conclude that the names *Neusticurus juruazensis* and *Ptychoglossus brevifrontalis* were mistakenly switched by Pellegrino et al. (2001) in the 12S and 16S data sets. With the addition of our specimen of *P. brevifrontalis*, this was obvious when examining the individual gene trees for 12S and 16S; *P. brevifrontalis* of Pellegrino et al. (2001) (GenBank nos. AF420757, AF420697) formed a clade with *N. eupleopus* and *N. strangulatus*, whereas *N. juruazensis* (GenBank nos. AF420758, AF420704) was virtually identical to our *P. brevifrontalis*. After examining the ND4 and c-mos trees, it was evident that these taxa were inadvertently switched in the 12S and 16S data sets. We swapped the names for these taxa for all of our reconstructions.

An additional problem stems from the inferred phylogenetic placement of *Bachia dorbignyi* (GenBank no. AF420861) and *Arthrosaura reticulata* (GenBank no. AF420841) in the c-mos gene partition. These two species had identical sequences and formed a close sister relationship with *Cercosaura eigenmanni*. This relationship can be seen in the squamate phylogeny reconstruction of Harris (2003), which used these same sequences. They differed from *C. eigenmanni* by one base substitution and appeared to be distantly related to other species of *Bachia* (47 and 51 base substitutions) and from the other three members of the Eupleopini, for which c-mos sequence was available by 52 to 89 base differences (*Colobosauroides cearensis*, 77; *Eupleopus gaudichaudii*, 52; *Leposoma oswaldoi*, 89). In all other gene partitions *Bachia dorbignyi* is closely related to its congeners and *Arthrosaura reticulata* is closely related to other eupleopiines. These phylogenetic positions are further supported by morphological features such as leg reduction in all species of *Bachia*. We cannot say with certainty what this sequence represents but it appears most similar to *Cercosaura eigenmanni*. Because we could not determine the true origin of the sequences, we excluded these two c-mos sequences from all final analyses.