Phylogenetic Systematics and Biogeography of Hummingbirds: Bayesian and Maximum Likelihood Analyses of Partitioned Data and Selection of an Appropriate Partitioning Strategy

JIMMY A. MCGUIRE,[1] CHRISTOPHER C. WITT,[2] DOUGLAS L. ALTSHULER,[3] AND J. V. REMSEN, JR.[4]

[1]*Museum of Vertebrate Zoology, 3101 Valley Life Sciences Building, University of California, Berkeley, CA 94720-3160, USA;*
E-mail: mcguirej@berkeley.edu (J.A.M.)
[2]*Museum of Southwestern Biology and Department of Biology, University of New Mexico, Albuquerque, NM 87131-0001, USA*
[3]*Department of Biology, University of California, Riverside, 900 University Avenue, Riverside, CA 92521-0427, USA; E-mail: douga@ucr.edu*
[4]*LSU Museum of Natural Science, 119 Foster Hall, Louisiana State University, Baton Rouge, LA 70803, USA; E-mail: najames@lsu.edu*

Abstract.—Hummingbirds are an important model system in avian biology, but to date the group has been the subject of remarkably few phylogenetic investigations. Here we present partitioned Bayesian and maximum likelihood phylogenetic analyses for 151 of approximately 330 species of hummingbirds and 12 outgroup taxa based on two protein-coding mitochondrial genes (ND2 and ND4), flanking tRNAs, and two nuclear introns (AK1 and BFib). We analyzed these data under several partitioning strategies ranging between unpartitioned and a maximum of nine partitions. In order to select a statistically justified partitioning strategy following partitioned Bayesian analysis, we considered four alternative criteria including Bayes factors, modified versions of the Akaike information criterion for small sample sizes (AIC$_c$), Bayesian information criterion (BIC), and a decision-theoretic methodology (DT). Following partitioned maximum likelihood analyses, we selected a best-fitting strategy using hierarchical likelihood ratio tests (hLRTS), the conventional AICc, BIC, and DT, concluding that the most stringent criterion, the performance-based DT, was the most appropriate methodology for selecting amongst partitioning strategies. In the context of our well-resolved and well-supported phylogenetic estimate, we consider the historical biogeography of hummingbirds using ancestral state reconstructions of (1) primary geographic region of occurrence (i.e., South America, Central America, North America, Greater Antilles, Lesser Antilles), (2) Andean or non-Andean geographic distribution, and (3) minimum elevational occurrence. These analyses indicate that the basal hummingbird assemblages originated in the lowlands of South America, that most of the principle clades of hummingbirds (all but Mountain Gems and possibly Bees) originated on this continent, and that there have been many (at least 30) independent invasions of other primary landmasses, especially Central America. [Akaike information criterion; ancestral state reconstruction; Bayesian information criterion; Bayes factors; biogeography; branch-length priors; decision theory; hierarchical likelihood-ratio test.]

Hummingbirds (Apodiformes: Trochilidae) have undergone a dramatic adaptive radiation and represent the second largest avian family in the New World, with ~331 described species in ~104 genera (Dickinson, 2003). Furthermore, they are ecologically diverse, with most hummingbird communities characterized by several sympatric species, and the richest sites inhabited by 25 or more taxa (for examples, see Robbins et al., 1987; Parker et al., 1985; Stiles and Levy, 1994; Krömer and Kessler, 2006; Dziedzioch et al., 2003). Because they exhibit substantial species richness and ecological diversity, approach the upper limits for vertebrate metabolism and physiological performance (Suarez et al., 1991; Chai and Dudley, 1995, 1996; Suarez, 1998), have coevolved extensively with many plant taxa (Grant and Grant, 1968; Temeles and Kress, 2003), and are relatively easy to manipulate both in the field and in the laboratory (Bartholomew and Lighton, 1986; Altshuler and Dudley, 2002), they represent an important avian model system. However, the great potential for this group as a model for comparative biology has not been fully realized, in part because it has not translated into phylogenetic effort. Here we provide a densely sampled, multilocus molecular phylogenetic estimate for 151 hummingbird species spanning the full extent of hummingbird higher-level diversity. This study builds upon a 75-taxon phylogenetic analysis published recently (Altshuler et al., 2004), a study focused on flight performance evolution that consequently did not explore in substantial detail the implications of the included phylogenetic estimate.

In the context of our phylogenetic estimate for hummingbirds, we evaluate several aspects of trochilid biogeography. Many aspects of hummingbird biogeography are of interest (see Chapman, 1917, 1926; Lack, 1973, 1976; Stiles, 1981; Bleiweiss, 1998; Rahbek and Graves, 2000), but we are primarily concerned here with the geographical sources of hummingbird faunas and the role that mountains have played in hummingbird diversification. We reconstruct ancestral areas of hummingbird occurrence to test the hypotheses that (1) hummingbirds originated in South America, (2) the South American Bee hummingbird clade reinvaded South America from temperate Central and North America, and (3) dispersal out of South America had a major influence on the assembly of hummingbird communities (Bleiweiss, 1998). Second, we reconstruct minimum elevational occurrences for every node on the hummingbird tree, in order to test the hypotheses that hummingbirds (1) originated in the lowlands and subsequently diversified extensively in upland habitats, and (2) invaded upland habitats unidirectionally from the lowlands. Finally, we evaluate Andean versus non-Andean occurrence to test the hypothesis that hummingbirds invaded the Andes one or a few times and radiated in situ, as opposed to hummingbirds having repeatedly invaded this particularly hummingbird-rich region.

Partitioned Phylogenetic Analyses

Partitioned phylogenetic analyses have generally been undertaken in a Bayesian framework because, until recently, mixed-model searching has not been possible using maximum likelihood (but see Stamatakis, 2006). Among those partitioned Bayesian studies that evaluated whether partitioning was statistically justified, most have relied on Bayes factors (e.g., Nylander et al., 2004; Brandley et al., 2005; Smith et al. 2005; Wiens et al. 2005; Moyle et al. 2006; but see Castoe et al., 2005, and Castoe and Parkinson, 2006). However, application of Bayes factors for phylogenetic model selection has been criticized because (1) they might not be sufficiently stringent in the phylogenetic context because the penalty imposed for additional parameters via the prior terms is insufficient (Pagel and Meade, 2004; Sullivan and Joyce, 2005); and (2) they are typically calculated from harmonic mean log-likelihoods, which are potentially biased in favor of more parameterized models (Lartillot and Philippe, 2006).

It is widely appreciated that application of underparameterized models in phylogenetic inference can lead to long-branch attraction problems and consequent phylogenetic error (Gaut and Lewis, 1995; Huelsenbeck, 1995; Sullivan and Swofford, 1997). The potential costs of overparameterization (including overpartitioning) are less intuitive, with parameter nonidentifyability, increased variance, the possibility of generating improper posterior distributions, and undue influence of the priors being the primary concerns (Chang, 1996; Rannala, 2002; Steel, 2005). In practice, many researchers employing mixed-model phylogenetic analysis have simply selected the most partitioned model available for their data sets without providing statistical justification (i.e., 15 of 16 studies published in this journal in 2006), suggesting that they are more concerned about the well-documented underparameterization problem than with overparameterization. Although we are sympathetic to the view that it is better to proceed with an overparameterized than underparameterized model, this does not validate ignoring the potential repercussions of estimating more parameters than are justified by the data (Steel, 2005). We follow Steel (2005), who advocates that the appropriate goal of model selection is to identify a model with sufficient (i.e., the minimum number of) parameters to "capture the key features of the data, including historical signal." Because application of standard Bayes factors using Kass and Raftery's (1995) conventions routinely results in selection of the most partitioned model under consideration (Sullivan and Joyce, 2005), we consider alternative model-selection criteria that impose explicit parameterization penalties, including a recently developed decision-theoretic approach originally implemented to evaluate alternative substitution models that we adapt to consider alternative partitioning strategies (Minin et al. 2003; Abdo et al. 2004; Sullivan and Joyce, 2005).

MATERIALS AND METHODS

Taxonomic Sampling and DNA Sequence Data

DNA sequences were obtained for 151 species of hummingbirds representing 73 of the approximately 104 currently recognized trochilid genera (online Appendix 1; available at www.systematicbiology.org). Our selection of outgroup taxa was based on the growing consensus that Apodiformes (hummingbirds and swifts) and Caprimulgiformes (nightjars and their relatives) form a monophyletic group (see Cracraft et al., 2004). We included 12 outgroup species spanning three avian orders including five swifts (Apodiformes, Apodidae): *Aerodramus salangana, A. vanikorensis, Aeronautes saxatalis, Chaetura pelagica,* and *Streptoprocne zonaris*; one treeswift (Apodiformes, Hemiprocnidae): *Hemiprocne mystacea*; five caprimulgiform birds including representatives of Aegothelidae (*Aegotheles insignis*), Caprimulgidae (*Caprimulgus longirostris*), Nyctibiidae (*Nyctibius bracteatus*), Podargidae (*Podargus strigoides*), and Steatornithidae (*Steatornis caripensis*); and one galliform bird (Phasianidae): *Coturnix chinensis*. For ingroup taxa and most of our outgroup species, we collected 4049 alignment positions of DNA sequence data representing four genes (two nuclear and two mitochondrial plus flanking tRNAs). The nuclear genes include intron 7 of beta fibrinogen (BFib) and intron 5 of the adenylate kinase gene (AK1). The mitochondrial gene sequences include the complete NADH dehydrogenase subunit 2 (ND2), approximately half of NADH dehydrogenase subunit 4 (ND4), and tRNAs flanking each of these protein-coding genes. The BFib sequences comprise 1406 aligned base pairs (bp), but this includes fairly extensive indels and most raw sequences were approximately 1100 bp in length. Likewise, the aligned AK1 sequences include 657 bp, but typical raw sequences were approximately 550 bp in length. The ND2 fragment comprises 1041 aligned base pairs, and the partial ND4 gene and its flanking tRNA sequences comprise 695 and 217 bp, respectively.

Several BFib sequences (for the outgroup species *A. vanikorensis, S. zonaris, P. strigoides, A. insignis, C. longirostris, S. caripensis,* and *H. mystacea*) were provided by the Early Bird Tree of Life project. The AK1 sequence for *A. insignis*, and the ND2 and ND4 sequences for *Coturnix chinensis* were obtained from GenBank.

DNA Sequence Data Collection and Alignment

DNA was isolated using standard phenol-chloroform extraction, a sodium-chloride extraction protocol, or using Qiagen DNeasy extraction kits following standard protocols. Amplification of target sequences was performed using the polymerase chain reaction with the primer sets identified in Table 1. Amplified PCR products were gel-purified on sephadex columns or purified using shrimp phosphatase and exonuclease (exoSAPit). A subset of nuclear gene sequences was cloned prior to

TABLE 1. Primers employed in this study.

Gene	Primer name	Primer sequence (5′-3′)	Source
BFib	FIB-B17U	GGAGAAAACAGGACAATGACAATTCAC	Prychitko and Moore, 1997
BFib	FIB-B17L	TCCCCAGTAGTATCTGCCATTAGGGTT	Prychitko and Moore, 1997
BFib	FIB-B17U2	CATCCATGCAGTTCTGGCAATTC	Prychitko and Moore, 1997
BFib	FIB-B17L2	TGGGAGGTGAAGCAGCTAAGAAAAACAA	Prychitko and Moore, 1997
BFib	FIB-B17U-long	GGAGAAAACAGGACAATGACAATTCACAATGG	This study
BFib	FIB-B17L-long	TCCCCAGTAGTATCTGCCATTAGGGTTGGC	This study
AK1	AK5b+	ATTGACGGCTACCCTCGCGAGGTG	Shapiro and Dumbacher, 2001
AK1	AK6c-	CACCCGCCCGCTGGTCTCTCC	Shapiro and Dumbacher, 2001
AK1	AK5b-extended	ATTGACGGCTACCCTCGCGAGGTGAAACAG	This study
AK1	AK6c-extended	CACCCGCCCGCTGGTCTCTCCTCG	This study
AK1	AK5b-inset	GGCTACCCTCGCGAGGTGAAACAG	This study
AK1	AK6c-inset	TGGTCTCTCCTCGCTTCAG	This study
ND2	H6313	CTCTTATTTAAGGCTTTGAAGGC	Sorenson et al., 1999
ND2	L5219	CCCATACCCCGAAAATGATG	Sorenson et al., 1999
ND2	H5766	GGATGAGAAGGCTAGGATTTTKCG	Sorenson et al., 1999
ND2	L5758	GGCTGAATRGGMCTNAAYCARAC	Sorenson et al., 1999
ND4	ND4	CACCTATGACTACCAAAAGCTCATGTAGAAGC	Arevalo et al., 1994
ND4	LEU	CATTACTTTTACTTGGATTTGCACCA	Arevalo et al., 1994

sequencing using TopoTA cloning kits because sequence length polymorphisms prevented us from obtaining clean sequences. However, most purified PCR products were sequenced directly using Big Dye Terminator sequencing reaction mix following manufacturer's protocols (Applied Biosystems). Cycle-sequencing products were visualized on either an ABI 377, ABI 3700, or ABI 3730 automated sequencer. All new DNA sequences generated for this study were deposited in GenBank (accession numbers EU042190-EU042594).

DNA alignments of the intron sequences were performed initially using the alignment program Muscle version 3.6 (Edgar, 2004), using default settings and further adjusted manually. Mitochondrial sequences were aligned manually, with tRNA sequence alignments adjusted to account for secondary structure (Kumazawa and Nishida, 1993; Macey and Verma, 1997). MacClade 4.06 (Maddison and Maddison, 2003) was used to verify that the ND2 and ND4 protein-coding gene sequences remained in frame throughout their lengths. The AK1, BFib, and tRNA sequences include a number of indels, some of which could not be aligned with confidence, and these regions were consequently excluded from further analysis. Heterozygous sites were scored using ambiguity codes and base positions contributing to length polymorphisms were excluded from analysis. The complete alignment is available on the TreeBASE website (study accession number S1825, matrix accession number M3354).

Phylogenetic Software and Parallel Computing

We performed maximum likelihood (ML) and Bayesian phylogenetic analyses using three programs: PAUP 4.0 (Swofford, 2003), RAxML 2.1.2 (Stamatakis, 2006), and a parallel version of MrBayes (version 3.1.1; Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003; Altekar et al., 2004) either compiled for analysis on a 15-node Linux Cluster or run in parallel on two dual-processor G4 Macintosh computers using Pooch interface software version 1.6 (http://daugerresearch.com/pooch/top.html). Model

selection and tree visualization were undertaken using PAUP. We performed ML analyses and associated nonparametric bootstrap analyses using RAxML, whereas Bayesian analyses were undertaken using MrBayes.

Model Selection

We identified best-fitting models for each data partition (see below) using the Akaike information criterion (AIC; Akaike, 1973; see Posada and Buckley, 2004, for justification), as implemented in ModelTest 3.06 (Posada and Crandall, 1998). When running ModelTest, the initial tree was drawn arbitrarily (tree 1) from the set of equally parsimonious trees obtained with the complete data. Because MrBayes 3.1.1 only implements 1, 2, and 6 substitution rate models, it was often not possible to implement the preferred model as selected by the AIC, and we were consequently forced to choose from a set of underparameterized or overparameterized models. In these situations, we selected the nearest overparameterized model because we are more concerned about the possible negative consequences of underparameterization (underestimated branch lengths and consequent long-branch attraction problems; Felsenstein, 1978; Huelsenbeck and Hillis, 1993; Gaut and Lewis, 1995; Sullivan and Swofford, 1997, 2001; Swofford et al., 2001; Anderson and Swofford, 2004) than we are about overparameterization in this context (potential parameter nonidentifiability). We base this assessment primarily on several simulation studies that found little cost associated with substitution model overparameterization, at least within the framework of the general time-reversible family of models (Lemmon and Moriarty, 2004; Huelsenbeck and Rannala 2004; see also Rannala, 2002). The models implemented in our Bayesian phylogenetic analyses are listed in Table 2. RAxML only implements the GTR+Γ model and is therefore applied to all partitions in ML analyses.

Data Partitions

Although our five gene-sequence data sets (corresponding to ND2, ND4, tRNAs, BFib, and AK1) were

TABLE 2. Alternative data partitions for phylogenetic analyses and DNA substitution models applied to each (Bayesian analyses only, RAxML only applies GTR+Γ for any partition). Abbreviations are defined as follows: mtDNA (all mitochondrial nucleotide sites), nuc (all nuclear gene nucleotide sites), mtpos1–3 (first, second, and third codon positions, respectively, for the ND2 and ND4 mitochondrial genes), ND2pos1–3 (first, second, and third codon positions for the ND2 mitochondrial gene), ND4pos1–3 (first, second and third codon positions for the ND4 mitochondrial gene), tRNA (transfer RNA genes flanking the ND2 and ND4 mitochondrial genes), AK1 (the entire AK1 intron fragment), BFib (the entire BFib intron fragment). In 4 and 5A partition analyses, tRNA sites were included in the mtpos2 partition. In 6B partition analyses, tRNA sites were included in the ND2pos2 partition.

Unpartitioned: (All nucleotide positions: GTR+I+Γ)

2 Partitions: (mtDNA, nuc: GTR+I+Γ)

4 Partitions: (mtpos1, mtpos2, mtpos3, nuc: GTR+I+Γ)

5 Partitions A: (AK1, BFib: GTR+Γ), (mtpos1, mtpos2, mtpos3: GTR+I+Γ)

5 Partitions B: (tRNA, mtpos1, mtpos2, mtpos3, nuc: GTR+I+Γ)

6 Partitions A: (tRNA, mtpos1, mtpos3, ND2pos2, ND4pos2, nuc: GTR+I+Γ)

6 Partitions B: (AK1, BFib: GTR+Γ), (mtpos1, mtpos3, ND2pos2, ND4pos2: GTR+I+Γ)

8 Partitions: (tRNA, ND2pos1, ND2pos2, ND2pos3, ND4pos1, ND4pos2, ND4pos3, nuc: GTR+I+Γ)

9 Partitions: (tRNA, ND2pos1, ND2pos2, ND2pos3, ND4pos1, ND4pos2, ND4pos3: GTR+I+Γ), (BFib, AK1: GTR+Γ)

---

initially analyzed individually (with individual analyses of the ND2 and ND4 genes partitioned by codon position), our ultimate goal was to analyze the data jointly. Because our combined data set is comprised of two protein-coding mitochondrial genes, tRNA sequences, and two nuclear introns, we suspected that application of a single nucleotide substitution model was unlikely to provide a particularly good fit to the data (e.g., Nylander et al., 2004: Brandley et al., 2005). Therefore, we undertook a battery of partitioned analyses implementing separate nucleotide substitution models for subsets of the data more likely to have experienced similar evolutionary processes. We evaluated nine distinct partitioning regimes ranging from unpartitioned to a maximum of nine partitions (Table 2). The unpartitioned analyses treated all sequence data under a common model of DNA substitution. Nine-partition analyses included separate substitution models for the tRNA sequences, the BFib and AK1 introns, and each codon position within the ND2 and ND4 mitochondrial genes.

Bayesian Phylogenetic Analyses

Markov chain Monte Carlo analyses were undertaken using MrBayes 3.1.1. Each analysis was allowed to run for 10 million generations sampling from the chain every 1000 generations. Following termination of these independent analyses, we first identified and discarded those sample points collected prior to convergence of the chains. We employed two strategies to confirm that the chains had achieved stationarity. First, we evaluated "burn-in" plots by plotting log-likelihood scores, tree lengths, and all model parameter values against generation number using the program Tracer (version 1.3) or the statistics program Statview version 5.0 (when it was nec-

essary to obtain a more detailed view of the burn-in plot). Second, we also evaluated convergence using the on-line application AWTY (Wilgenbush et al., 2004) by performing analyses of posterior probability clade-support values using the cumulative function to verify that these values were stable across all post-burn-in generations within each analysis. Posterior probability values are not expected to vary directionally (either upward or downward) over time once the Markov chain has achieved stationarity and substantial deviations from an equilibrium value over time would suggest that the chain had not yet converged. After determining chain convergence, which generally occurred within the first 2 to 5 million generations of each analysis, we followed a conservative approach by discarding all samples obtained during the first 8 million generations as "burn-in." We then generated consensus phylograms with mean branch-length estimates, 50% majority-rule consensus trees with posterior probability values for each node, credible sets of trees, and parameter estimates in the context of the final 2 million generations obtained during each analysis.

We used two strategies to reduce the opportunity for our analyses to become trapped on local optima. First, we employed Metropolis-coupled Markov chain Monte Carlo with each analysis including a cold chain and three incrementally heated chains (the default in MrBayes 3.1). When a chain is heated, the magnitude of difference between two likelihood scores is compressed relative to the difference that would be observed with unheated chains. This makes it easier for heated chains to cross deep valleys, thus allowing a more efficient exploration of the likelihood surface (Huelsenbeck and Ronquist, 2001). Preliminary analyses indicated that the default temperature regime for heating (T = 0.2) resulted in infrequent or complete absence of swapping of states between the heated and cold chains. We found that T = 0.02 resulted in state swap frequencies of 20% to 80% (an appropriate target value according to the MrBayes 3.1.1 manual and J. P. Huelsenbeck, personal communication) and consequently used this temperature regime for all subsequent analyses. Second, we compared recovered topologies and associated posterior probability values obtained across four independent analyses undertaken for each data set (four separate analyses with Nruns = 1), each of which was initiated from random starting points. These comparisons were undertaken using the compare function in AWTY. Once we discarded the pre-burn-in samples and verified that the four independent runs had converged on similar stationary distributions, we consolidated the data from the four analyses for final processing. Thus, the Markov chains for each partitioning strategy were effectively allowed to run for 40 million generations.

Although we ultimately selected the best-fitting partitioning strategy using a decision-theoretic methodology (see below), we also used the "compare" function in AWTY to evaluate the degree of agreement between topologies and branch-support values obtained under the nine alternative partitioning strategies. Our motivation for this analysis was to determine whether inferred

hummingbird branching arrangements and support values were robust to alternative partitioning strategies.

All partitioned analyses accommodated among-partition rate variation (APRV) by using the "prset ratepr=variable" option in MrBayes. We chose to employ this option following the recommendations of Marshall et al. (2006) after finding that our partitioned analyses that explicitly accommodated APRV resulted in trees with two- to threefold shorter branch length estimates and greater harmonic mean log likelihoods (HML_i) than did equivalent analyses that did not accommodate APRV. We also performed preliminary analyses under several branch-length priors based on the finding of Marshall et al. (2006) that use of the default prior led to poor branch-length estimation with their empirical data set. We undertook separate analyses with the eight-partition and nine-partition data sets using the default exponential prior on branch-length mean = 10 ("brlens=unconstrained:Exp[10]"), as well as exponential distributions with means of 2, 50, and 100. In these preliminary analyses, we found that the branch-length exponential prior with a mean of 50 substantially outperformed other prior values in the eight-partition and nine-partition analyses both in terms of a significantly higher HML_is for topologies in the posterior distribution and in better convergence behavior and mixing (analyses using alternative priors did not converge to a stationary distribution, which we evaluated by examining likelihood scores and model parameter values plotted against generations). However, when we undertook analyses with six or fewer partitions, the default exponential branch-length prior with mean = 10 either outperformed or performed equivalently to the mean = 50 exponential distribution, further supporting the notion that multiple branch-length priors should be considered when undertaking Bayesian analyses. Our final analyses, with eight partitions and nine partitions, respectively, employed the branch-length exponential prior with mean = 50, whereas all other analyses employed the default exponential branch-length prior with mean = 10.

Maximum Likelihood Analyses

Partitioned and unpartitioned ML analyses were undertaken using the parallel version of RAxML (RAxML HPC-MPI; Stamatakis, 2006). For each partitioning regime, two analyses were undertaken. The goal of the first set of analyses was to obtain the maximum-likelihood estimate (for likelihood ratio tests; see below) assuming the Bayesian nine-partition consensus topology under each partitioning regime. These analyses employed the GTRGAMMA substitution model, which will provide comparable likelihood estimates (unlike the more computationally efficient GTRCAT approximation [Stamatakis, 2006]). Importantly, with partitioned ML analyses in the current implementation of RAxML, a single overall rate of evolution is optimized across all partitions. This probably results in branch-length estimates that are not as accurate as those that would be obtained using MrBayes because MrBayes allows for the

application of an evolutionary rate multiplier to individual partitions. This rate multiplier allows evolutionary rates to vary across partitions under a flat dirichlet prior, and accommodating among-partition rate variation has been shown to result in better branch-length estimation (Marshall et al., 2006). In the second set of ML analyses for each partitioning regime, we searched for the ML topology, using 200-replicate bootstrap analyses to estimate nodal support. Given the computational challenges associated with ML nonparametric bootstrapping, this second set of analyses employed the GTRCAT approximation.

Identifying an Appropriate Partitioning Strategy

We evaluated several decision criteria for selecting alternative partitioning strategies. We applied standard Bayes factors (Kass and Raftery, 1995; Nylander et al., 2004), only accepting Bayes factors greater than 10 as evidence in support of a more partitioned model (2 ln Bayes factors >10 are considered to be "very strong" according to the Kass and Raftery's [1995] conventions). In addition, we evaluated alternative partitioning strategies using modifications of the AIC_c, BIC (Schwarz, 1978), and a decision-theoretic (DT) approach (Minin et al., 2003; Abdo et al., 2004; Sullivan and Joyce, 2005). The AIC_c, BIC, and DT are normally applied in a maximum likelihood framework, but we have also used them in a Bayesian framework under the assumption that the HML_is under alternative partitioning strategies exhibit a similar relationship to one another as would maximum likelihood estimates (see Castoe et al., 2005; Castoe and Parkinson, 2006, for a similar approach involving Akaike weights). When calculating the AIC_c, BIC, and DT for evaluation of Bayesian phylogenetic estimates, we substituted the HML_i for maximum likelihood values. We employed the following small sample-size modification of the AIC equation

$$AIC_c = -2\,HML_i + 2p_i + 2p_i(p_i + 1)/(n - p_i - 1)$$

in which p_i is the number of parameters in the model and n is the number of nucleotide sites. We used the following BIC equation

$$BIC = -2HML_i + (p_i + 2N - 3)\ln n$$

in which $2N - 3$ is the number of branch lengths to be estimated (see Minin et al., 2003). Because the number of branch lengths are constant across alternative partitioning strategies, this version of the BIC equation will return identical BIC estimates as would the standard BIC equation that imposes a parameterization penalty by simply taking the natural log of the number of substitution model parameters. The DT methodology involves calculation of a risk function based on the standard BIC, but with an additional penalty imposed if greater parameterization (in this case additional partitioning) results in greater variance in branch-length estimates (Minin et al., 2003; Abdo et al., 2004). The program

DT-modsel implements the DT approach for selection of nucleotide substitution models (script available at www.webpages.uidaho.edu/~jacks/DTModSel.html). In order to apply this approach to the selection of an appropriate partitioning strategy, we modifed the DT-modsel script to reflect the number alternative partitioning strategies (i.e., nine) that we wished to compare and adjusted the number of free parameters to reflect the number required for each partitioning regime. The available DT-modsel perl script requires that all models be compared in the context of a fixed topology. Therefore, to apply the procedure to our Bayesian analyses, the BIC component of the DT calculation was based on HML_is obtained under alternative partitioning strategies, but branch-lengths were recalculated on the nine-partition topology. We obtained fixed-topology branch-length estimates using MrBayes by disabling the program's branch-swapping functionality and enabling the node-slider (by resetting props; see MrBayes version 3.1 manual). The preferred partitioning strategy has the minimum observed AIC_c, BIC, or risk function (DT). The BIC imposes a heavier parameterization penalty than does the AIC and is therefore expected to favor less partitioned models, whereas the DT tends to be most stringent and is expected to select less partitioned models than either the AIC or BIC (Minin et al., 2003; Abdo et al., 2004).

We compared the partitioning strategies selected in a Bayesian framework (using Bayes factors, the AIC_c, BIC, and DT) with partitioning strategies selected using the hLRT, AIC_c, BIC, and DT in an ML framework. Applying these model-selection criteria required calculating maximum likelihoods under our nine alternative partitioning strategies using RAxML (see above). ML and Bayesian model-selection procedures were not strictly comparable because RAxML only implements a GTR+Γ model, whereas our Bayesian analyses employed a mixture of GTR+Γ and GTR+I+Γ models. Furthermore, all ML comparisons were in the context of the nine-partition Bayesian consensus tree, whereas Bayesian model-selection criteria were computed in the context of optimized trees for each partitioning strategy (except for DT).

In order to apply these model-selection criteria to either Bayesian or ML analyses, we need to calculate the number of parameters in each model. For Bayesian anlayses, the number of parameters is obtained by summing the total number of substitution rates (six per partition, corresponding to the GTR model), rate heterogeneity parameters (one per partition, corresponding to Γ), the number of invariant sites parameters (zero or one per partition, corresponding to I), and the number of among-partition rate multipliers (equal to the number of data partitions in the model). The number of parameters in partitioned ML and Bayesian analyses differ slightly, because ML analyses do not incorporate among-partition rate variation parameters, and have one fewer substitition rate parameter (because the substitution matrix is scaled against a fixed G-T rate of 1.0, whereas Bayesian analyses employ a Dirichlet prior that allows all rates in the substitution matrix to vary).

Ancestral State Reconstruction

To reconstruct the biogeographic history of hummingbirds, we utilized ancestral state reconstructions for three variables of interest: source landmass, Andean occurrence, and elevation. The source landmass analysis also could have been undertaken using a more explicit biogeographical methodology such as dispersal-vicariance analysis (DIVA: Ronquist, 1996), but we opted for an ancestral state reconstruction approach because it allowed us to utilize branch length information and to account for phylogenetic uncertainty by considering a posterior distribution of trees rather than a point estimate. Source landmass was scored as a five-state categorical character, with South America, Central America, North America, Greater Antilles, and Lesser Antilles being the five regions of interest. Geographic distribution data were primarily drawn from the AOU checklist (1998) and Schuchmann (1999). For this character, we performed ancestral state reconstructions for each node on our hummingbird tree using a likelihood method implemented in software written by J. P. Huelsenbeck (available from the authors upon request). The method calculates the probability of each of the (five) possible ancestral character states at each node conditional on the trees, total tree lengthes, relative branch length proportions, and tip values. We assume a Jukes-Cantor–like equal-rates model of character state transformation. We apply this simple model under the assumption that (1) all state changes are equally likely, and (2) there is unlikely to be sufficient information in a single character to allow for reliable estimation of the many parameters that would comprise a more complex model (such as the 20-rate nonreversible model that is the default for this character in BayesTraits; Pagel et al., 2004). Here, we took phylogenetic uncertainty into account by evaluating ML support for each nodal reconstruction across a set of 6000 trees drawn from the posterior distribution (see Lutzoni et al., 2001; Buschbom and Parker, 2006) obtained during our preferred partitioned Bayesian analysis (model 6B, see results for rationale). The ancestral state probabilities for a particular node were then taken as the means of the probabilities observed for that node across all trees. In order to account for uncertainty regarding appropriate branch length scaling (which can have a dramatic impact on ancestral state reconstructions), each tree in the posterior distribution was reevaluated 100 times under alternative tree length rescalings drawn from a gamma distribution. Thus, we ultimately calculated ML reconstructions for each node on 600,000 trees (100 permutations for each of 6000 trees in the posterior distribution). We attempted to further account for tree length scaling uncertainty by performing three separate analyses, one in which we set the mean tree length to be 13.0 (approximately equal to the empirically estimated treelength) with a variance of 5.0, another with a mean of 1.0 and a variance of 0.5, and a third with a mean tree length of 30.0 and a variance of 10.0. Longer tree lengths are expected to result in less certain ancestral state reconstructions because nodal values are less

TABLE 3. Percentages and numbers of alignable (included) base pairs, variable sites, and parsimony-informative sites for the gene fragments employed in this study.

	Included bps	% Variable sites (no. variable sites)	% Parsimony-informative sites (no. parsimony informative sites)
tRNA	210	0.486 (102)	0.333 (70)
ND2pos1	347	0.674 (234)	0.550 (191)
ND2pos2	347	0.412 (143)	0.300 (104)
ND2pos3	347	1.000 (347)	0.997 (346)
ND2	1041	0.695 (724)	0.616 (641)
ND4pos1	231	0.519 (120)	0.398 (92)
ND4pos2	232	0.259 (60)	0.147 (34)
ND4pos3	232	1.000 (232)	0.996 (231)
ND4	695	0.593 (412)	0.514 (357)
AK1	635	0.578 (367)	0.398 (253)
BFib	1245	0.643 (801)	0.455 (566)
Total	3826	0.629 (2406)	0.493 (1887)

likely to retain the values observed at their descendent tips.

Andean occurrence was scored as a binary categorical character, with ancestral state reconstructions undertaken using the maximum likelihood Mk1 Markov model implemented in Mesquite (Maddison and Maddison, 2005). Here, a taxon was scored as present in the Andes if more than half of its distributional range is included in one or more Andean distributional centers (Cracraft, 1985) or its elevational distribution extends above 1500 m in the main Andean Cordillera.

Elevation was scored as a continuous variable and analyzed using weighted squared-change parsimony in Mesquite (Maddison and Maddison, 2005). Weighted squared-change parsimony provides a maximum posterior probability and maximum-likelihood estimates for ancestral states under a Brownian model (Maddison, 1991; Schluter et al., 1997) and consequently takes branch-length information into consideration like other likelihood methods. We compiled three elevation variables (minimum, mid, and maximum elevational occurrence; data available upon request) from Parker et al. (1996).

RESULTS

Summaries of the number of nucleotide sites, percentage of variable sites, and number and percentage of parsimony-informative sites for each gene segment and for the complete matrix are provided in Table 3. Ranges of pairwise uncorrected sequence divergence values by gene region were as follows: tRNA (0–13.72%), ND4 (0–20.76%), ND2 (0.10%–23.27%), AK1 (0–10.47%), and BFib (0–9.46%).

Analyses of Individual Gene Segments

Bayesian phylogenetic analyses of the five individual gene regions (tRNA, ND2, ND4, BFib, and AKI) resulted in informative phylogenetic estimates for hummingbirds, although the degree of resolution and branch support varied widely by marker. The number of strongly supported ingroup nodes (with ≥0.95 posterior probability values) recovered from the four independent MCMC

runs per data set were as follows: tRNA (14 strongly supported nodes, 147 ingroup taxa), ND2 (101 nodes, 151 ingroup taxa), ND4 (70 nodes, 148 ingroup taxa), AK1 (47 nodes, 144 ingroup taxa), and BFib (83 nodes, 149 ingroup taxa). Despite this variance in degree of resolution and numbers of well-supported nodes, there was little significant disagreement among the markers. This was particularly true for the ND2, ND4, and tRNA data sets, which is not surprising given that these markers represent a single linkage group. The trees obtained for each gene region are provided as supplemental materials (available at www.systematicbiology.org).

Choosing an Optimal Partitioning Strategy

With the Bayesian analyses, Bayes factors, AIC$_c$, BIC, and DT model-selection criteria preferred three alternative partitioning strategies. Bayes factors and the AIC$_c$ prefer the fully partitioned (nine-partition) model, whereas the BIC selects the 8-partition model, and DT selects a six-partition model (model 6B; Table 4). In the ML framework, the hLRT and AIC$_c$ prefer the nine-partition model, whereas the BIC prefers a six-partition model (model 6B), and DT selects a five-partition model (model 5B; Table 4). For reasons outlined below (see Discussion), we conclude that model 6B (the model selected in a Bayesian framework by DT) is optimal for our Bayesian analyses, and that additional partitioning does not result in a sufficient increase in the HML$_i$s or maximum likelihoods to justify increased parameterization. The decision regarding which partitioning regime to employ is not critical with respect to our hummingbird phylogenetic estimate because the more-partitioned strategies considered here all return similar topologies and branch support values (see Figs. 1, 3, 4). However, we interpret this as support for the DT approach because its less partitioned model returns a similar topology and branch lengths as more partitioned models while estimating fewer parameters. Tree-length estimates vary only slightly across partitioning strategies, suggesting that much of the improvement in the likelihoods obtained with more extensive partitioning was probably associated with substitution model and base frequency parameter estimation (effectively nuisance parameters in this context) rather than with more critical topology and branch-length estimation. Although the topology and branch-length estimates obtained under the less partitioned model preferred by DT did not differ substantially from those obtained with more partitioned models (see Figs. 3, 4), we suspect that other taxonomic groups will be found for which selection of overpartitioned models will return relatively poor phylogenetic estimates.

Analyses of Combined Multilocus Data under Alternative Partitioning Strategies

Bayesian analysis under partitioning regime 6B resulted in a well-resolved and well-supported estimate of hummingbird phylogeny, with 123 of 151 ingroup nodes receiving posterior probability values ≥0.95 and nine additional nodes receiving posterior probabilities

TABLE 4. Data partition statistics including number of partitions, number of parameters per partitioning scheme, and model-selection criteria. Values in bold indicate the preferred partitioning strategy according to that model-selection criterion. Asterisks denote the model selected using the decision-theoretic approach. Complete descriptions of the partitions are provided in Table 2. In the matrix below, values above the diagonal are twice the difference in the HML_is for the alternative partitioning strategies, with values in bold highly significant according to Kass and Raftery (1995) conventions (2 ln Bayes factors >10). The numbers below the diagonal are twice the difference in the ML values under alternative partitioning strategies, with those in italics indicating non-nested models and those in bold significant according to the χ-square distribution ($\alpha = 0.05$).

Bayesian analyses

No. partitions	No. parameters	$-HML_i$	AIC_c	BIC
1	11	94,809.5	189,641.1	192,408.3
2	24	92,880.2	185,808.7	188,657.3
4	48	91,659.9	183,417.0	186,416.3
5A	58	91,604.8	183,327.3	186,389.1
5B	60	91,509.5	183,140.8	186,215.1
6A	72	91,544.2	183,235.1	186,384.1
6B	70	91,478.3	183,099.1	186,235.7*
8	96	91,338.7	182,874.1	**186,172.3**
9	108	**91,296.0**	**182,814.0**	186,186.6

ML analyses

No. partitions	No. parameters	$-L_i$	AIC_c	BIC
1	10	95,175.2	190,370.5	193,131.4
2	20	93,559.5	187,159.2	190,066.0
4	40	91,649.0	183,378.8	186,328.1
5A	48	91,602.7	183,302.6	186,301.9
5B	50	91,577.7	183,256.7	186,268.5*
6A	60	91,546.9	183,215.6	186,289.9
6B	58	91,542.4	183,202.5	**186,264.3**
8	80	91,468.7	183,100.7	186,299.5
9	90	**91,422.9**	**183,030.0**	186,290.9

	P9	P8	P6B	P6A	P5B	P5A	P4	P2	P1
P1	**7027.0**	**6941.6**	**6662.4**	**6530.6**	**6600.0**	**6409.4**	**6299.2**	**3858.6**	—
P2	**3168.4**	**3083.0**	**2803.8**	**2672.0**	**2741.4**	**2550.8**	**2440.6**	—	**3231.4**
P4	**727.8**	**642.4**	**363.2**	**231.4**	**300.8**	**110.2**	—	*3821.0*	*7052.4*
P5A	**617.6**	**532.2**	**253.0**	**253.0**	**190.6**	—	*92.6*	*3913.6*	*7145.0*
P5B	**427.0**	**341.6**	**62.4**	−69.4	—	*50.0*	**142.6**	*3963.6*	*7195.0*
P6A	**496.4**	**411.0**	**131.8**	—	**61.6**	*111.6*	**204.2**	*4025.2*	*7256.6*
P6B	**364.6**	**279.2**	—	*9.0*	*70.6*	**120.6**	*213.2*	*4034.2*	*7265.6*
P8	**85.4**	—	**147.4**	**156.4**	**218.0**	**268.0**	**360.6**	*4181.6*	*7413.0*
P9	—	**91.6**	**239.0**	**248.0**	**309.6**	**359.6**	**452.2**	*4273.2*	*7504.6*

between 0.90 and 0.95 (Fig. 2). Most of these strongly supported nodes (120 of 123) received posterior probabilities ≥0.95 under all nine alternative partitioning strategies (see Figs. 3, 4). Of the 124 strongly supported ingroup nodes, six received posterior probability values ≥0.95 in independent analyses of all five gene regions (i.e., tRNA, ND2, ND4, BFib, AK1; note that only 14 nodes were strongly supported by tRNA data alone), 23 were significantly supported in analyses of four gene regions, 20 were significantly supported in analyses of three gene regions, 37 were significantly supported in analyses of two gene regions, 28 were significantly supported in analyses of one gene region, and 9 were not significantly supported in analyses of individual gene regions but were nevertheless strongly supported in analyses of the combined data. Finally, 66 strongly supported nodes in the nine-partition analysis were significantly supported by at least two independent genetic loci.

The ML analysis under partitioning regime 5B (the model selected by DT in an ML framework) resulted in a topology highly similar to that recovered in the Bayesian analysis under regime 6B (146 of 151 ingroup nodes shared), with 115 ingroup nodes receiving bootstrap proportions equal to or greater than 0.70 (see supplemental materials at www.systematicbiology.org). Seven nodes that received posterior probability values ≥0.95 in the Bayesian analysis received bootstrap proportions between 0.63 and 0.68. One node (subtending *Campylopterus villaviscensio* and *C. hyperythrus*) was well supported in the ML analysis (BP = 0.71) but not in the Bayesian analysis (posterior probability = 0.85).

DISCUSSION

Given their remarkable diversity and the ease with which they can be manipulated both in the field and in the laboratory, we believe that hummingbirds have the potential to be the most important model system for avian comparative biology. However, to fulfill this potential, a densely sampled, well-resolved, and well-supported phylogenetic estimate for this group that can serve as the historical framework for evolutionary analyses is essential. The present study represents a step in this direction by providing a highly resolved and strongly supported multilocus phylogenetic estimate for a substantial component of hummingbird diversity (i.e., 151 of ~331 species and 73 of ~104 genera).
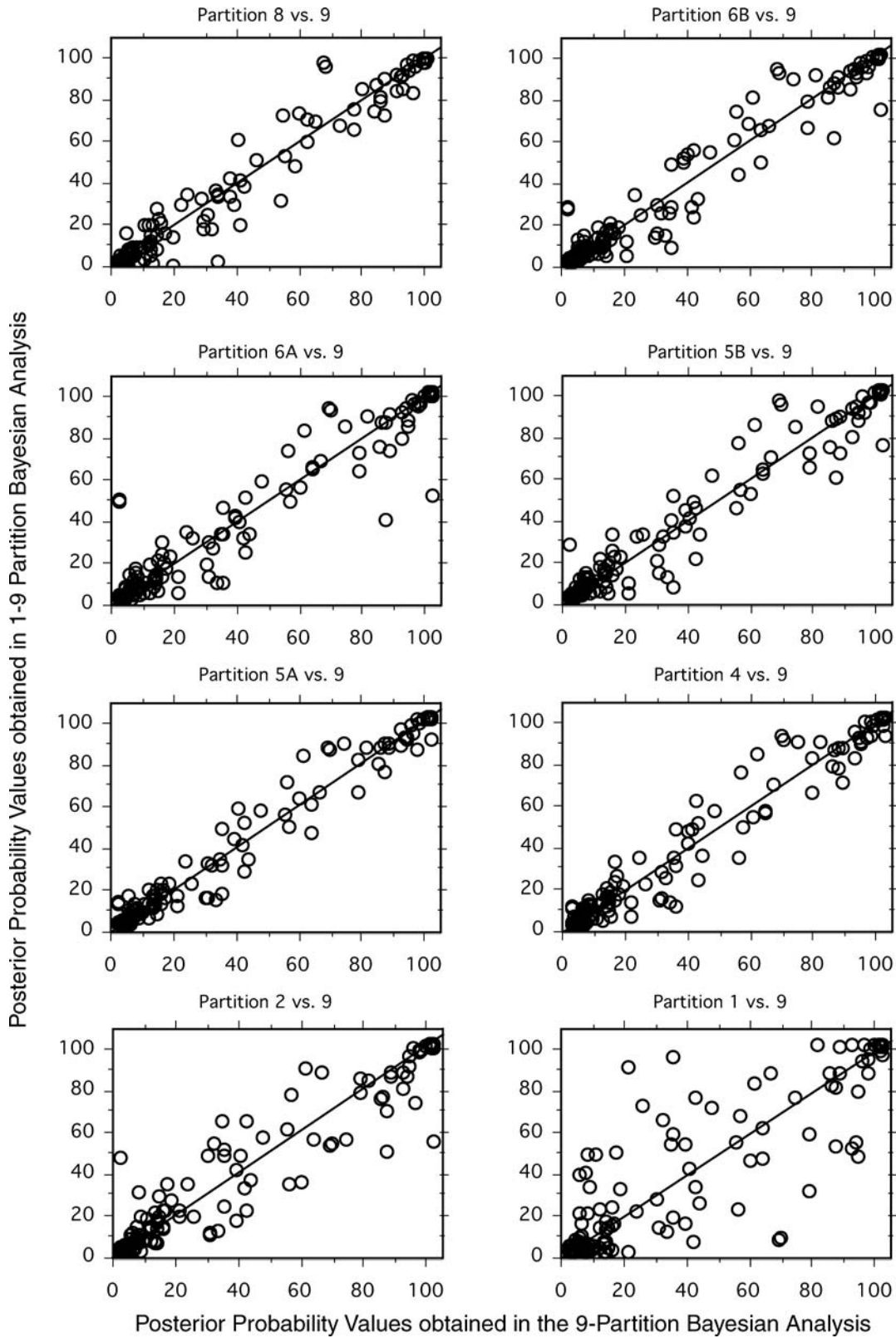
FIGURE 1. Plot of posterior probabilities from analyses under each partitioning strategy against posterior probabilities obtained in the nine-partition analysis.
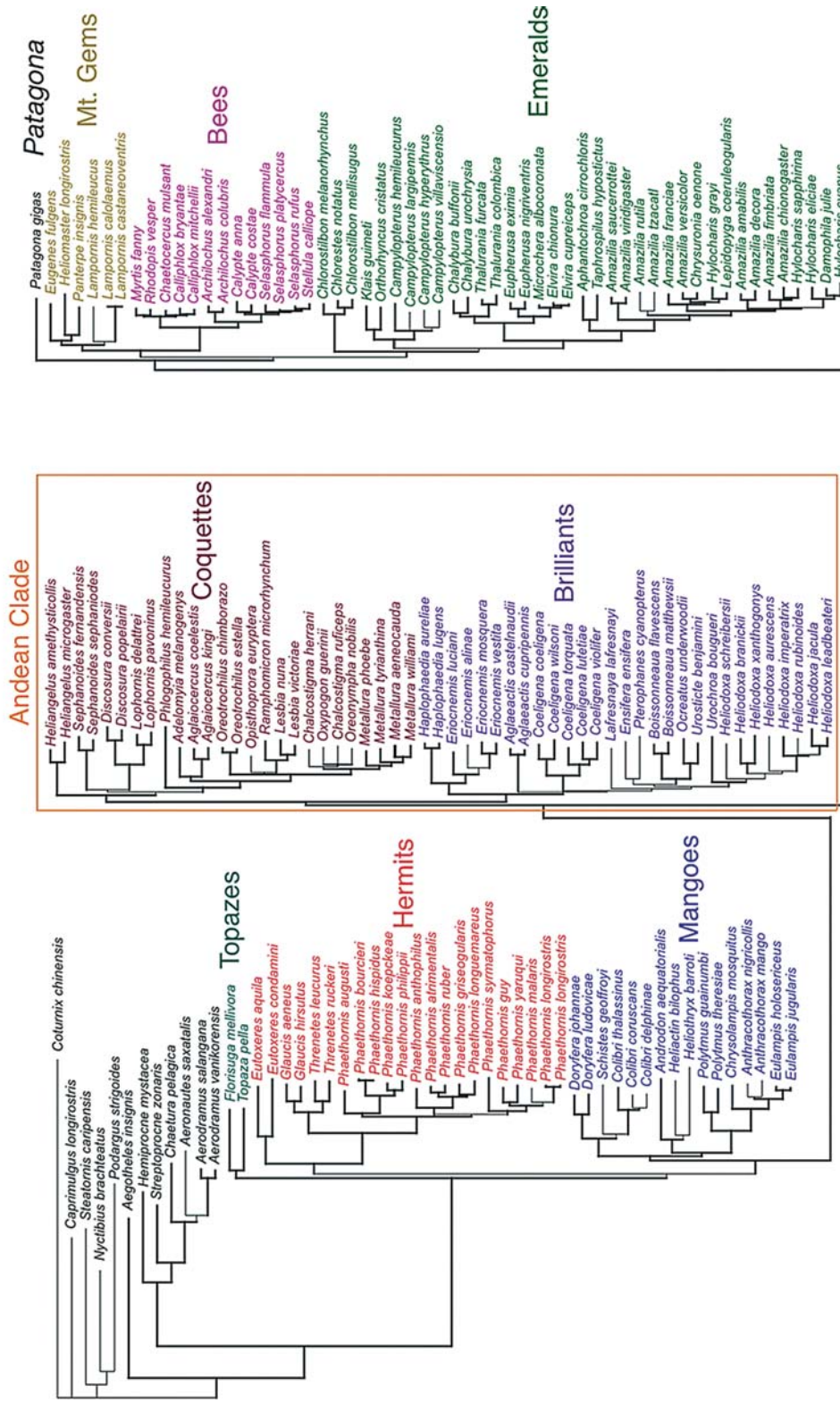
FIGURE 2. Phylogenetic estimate for hummingbirds based on 6B-partition analysis. The topology is the Bayesian consensus tree, and thicker branches have associated posterior probability values ≥0.95.
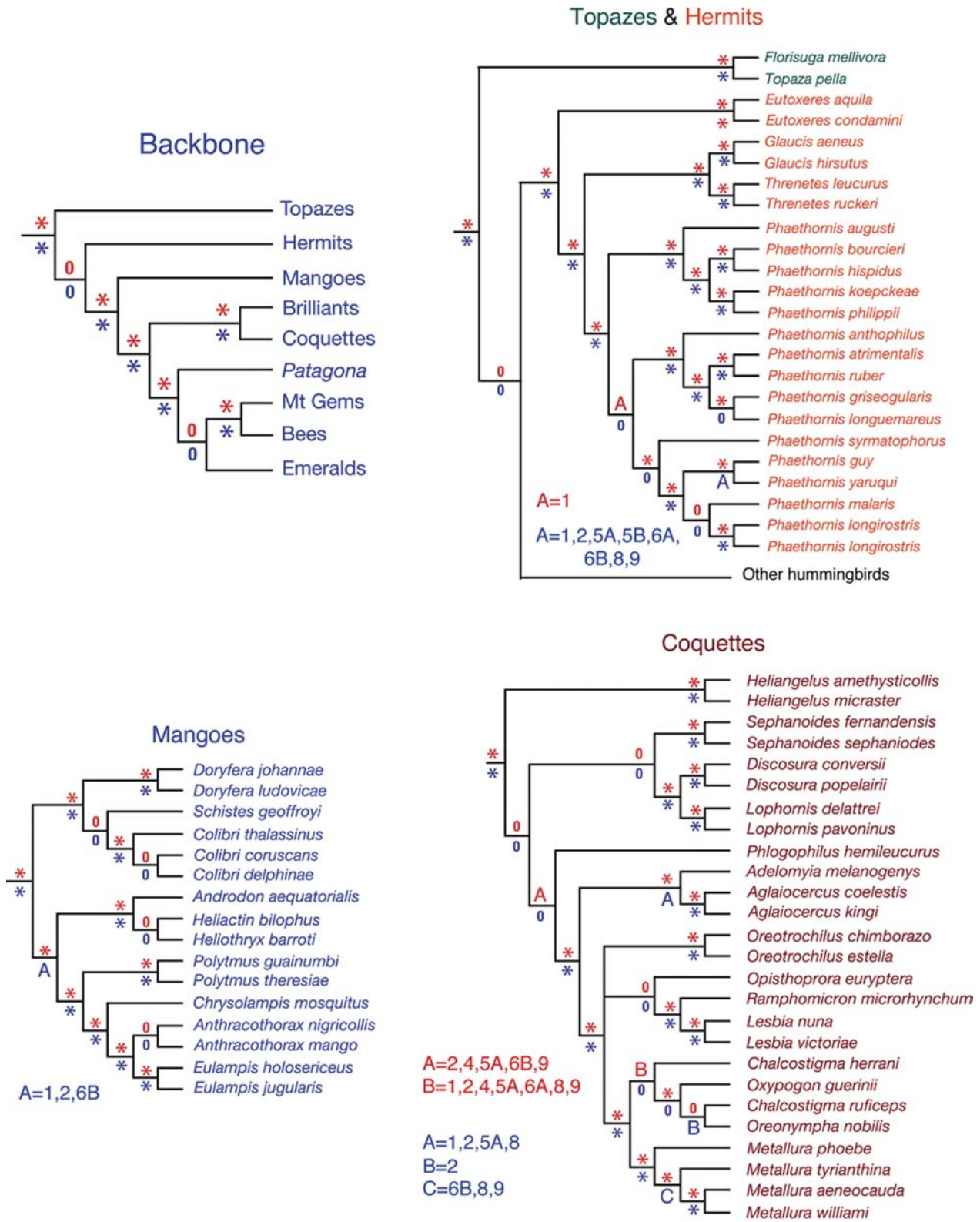
FIGURE 3. Phylogenetic estimates for the primary clades of hummingbirds, plus Topazes, Hermits, Mangoes, and Coquettes. The topologies are taken from the 6B-partition Bayesian consensus tree. All nodes indicated by a red asterisk above the branch received ≥0.95 posterior probability values under all nine partitioning strategies. Similarly, those nodes indicated by a blue asterisk below the branch received greater than 0.70 bootstrap proportions under all nine partitioning strategies. Nodes denoted by a red or blue "0" did not receive posterior probability values ≥0.95 (red) or bootstrap values ≥0.70 (blue) under any partitioning strategy. Those nodes denoted by a red or blue capital letter either received posterior probability values ≥0.95 (red) or bootstrap proportions ≥0.70 (blue) under those partitioning strategies indicated adjacent to the tree.
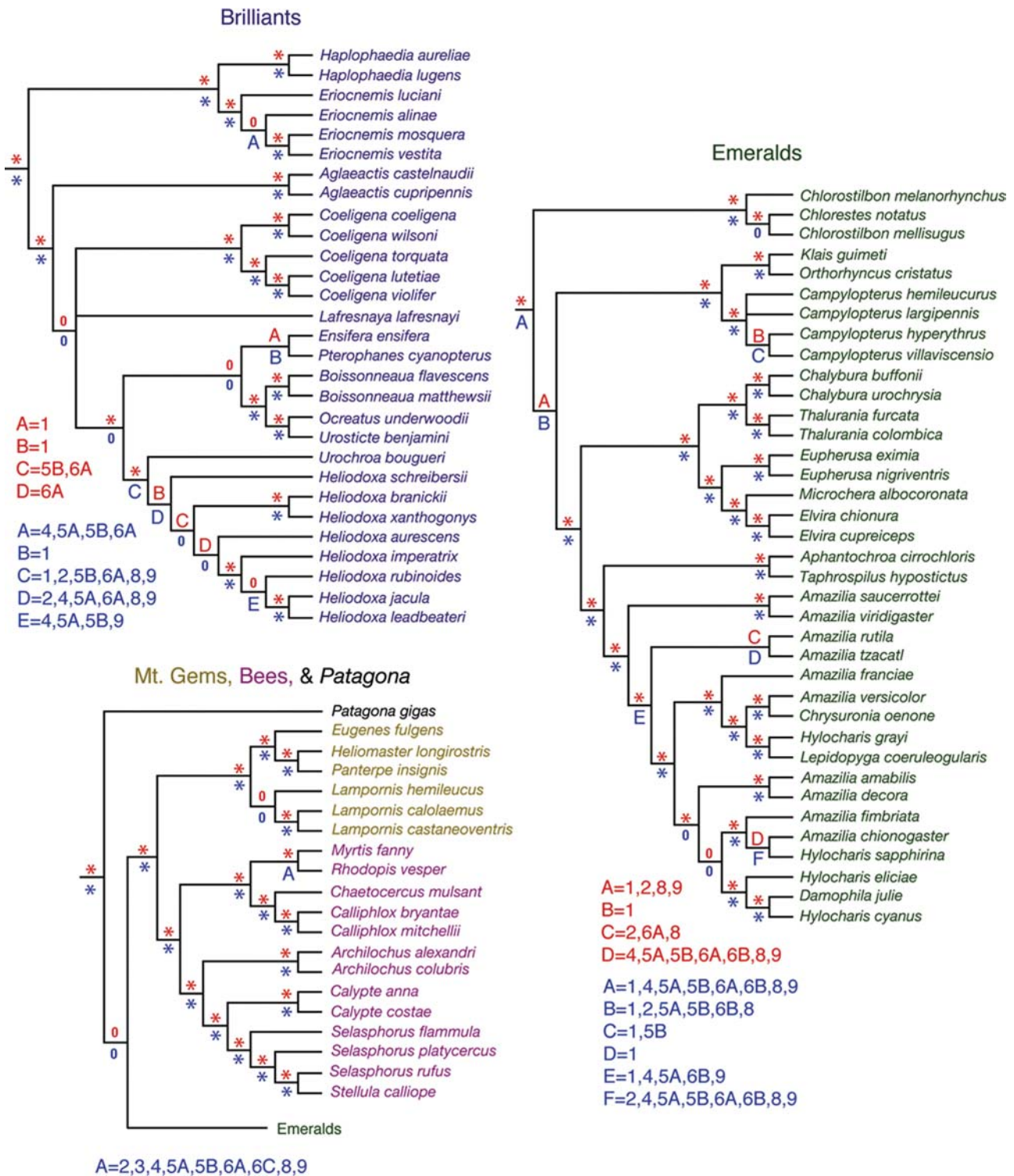
FIGURE 4. Phylogenetic estimates for Brilliants, Emeralds, Mountain Gems, Bees, and *Patagona*. Format is the same as for Figure 3.

Partitioned Phylogenetic Analyses of Multilocus Sequence Data

We considered nine alternative partitioning strategies in both Bayesian and ML frameworks and employed several alternative model-selection criteria to choose the best-fitting partitioning scheme. With our data, we found that the most-partitioned (nine-partition) model was selected by those criteria that impose relatively weak penalties for additional parameterizations (standard Bayes factors, AIC_c, and hLRT), whereas application of more

stringent criteria (BIC and DT) results in less than fully partitioned models being selected (models 8 and 6B, respectively). We note that the DT model-selection criterion has an important advantage relative to the other criteria we considered in that it takes into account relative performance in branch-length estimation, rather than just comparing the relative fit of alternative models (Minin et al., 2003, Abdo et al., 2004, Sullivan and Joyce, 2005). For example, more partitioned models could be selected by the non-performance-based criteria simply because they provide better estimation of nuisance parameters such as base frequencies and substitution rates. If better fitting of these parameters comes at the cost of inferior topology and/or branch-length estimation, then we would clearly be better off with a less-partitioned model. In the present study, we cannot presume to know which partitioning strategy results in the most reliable phylogenetic estimate, but we do know that the topology, branch lengths, and support values stabilize once the data are partitioned into four subsets corresponding to nuclear sequences and the three mitochondrial codon positions (with tRNAs lumped with mt second positions; see Fig. 1). We conclude that much of the improvement in ML and HML$_i$ estimates with greater partitioning is associated with greater precision in the estimation of nuisance parameters. That said, even the DT prefers models with five (ML) and six (Bayesian) partitions rather than four partitions, indicating that either the improvement in likelihood estimates associated with better fitting of nuisance parameters outweighs the cost function associated with poorer branch-length estimation or that there are subtle improvements in branch length estimation that justify one or two additional partitions.

Many published partitioned Bayesian phylogenetic studies have not bothered to test the statistical justifiability of their data partitioning strategies (i.e., 15 of 16 studies published in this journal in 2006), and those that have evaluated alternative partitioning schemes have relied primarily on Bayes factors (e.g., Nylander et al., 2004; Brandley et al., 2005; Smith et al., 2005; Wiens et al., 2005; Moyle et al., 2006; but see Castoe et al., 2005, and Castoe and Parkinson, 2006). With one exception (Moyle et al., 2006), these studies have selected the most partitioned model under consideration as best-fitting, causing Sullivan and Joyce (2005) to question the reliability of Bayes factors for partitioning model selection. Our results are consistent with the concerns expressed by Sullivan and Joyce (2005) in that Bayes factors (as well as the AIC$_c$) strongly support the most partitioned model despite the fact that similar topologies, branch lengths, and support values are obtained with fewer partitions. Although we are confident that the DT model-selection criterion provides a reliable approach, we note that Pagel and Meade (2004) have suggested simply increasing the stringency of the Bayes factor significance threshold. In their own study, they suggested a criterion of 2 ln Bayes factor >40, an arbitrary value that would result in the selection of model 5B with our data. In lieu of a controlled simulation study of this issue, we cannot reach a strong conclusion regarding which of the selection criteria considered here

is most appropriate. However, we do believe that the future of mixed-model phylogenetic analysis will include explicit testing of alternative partitioning strategies and more stringent parameterization penalties than are generally now employed.

Congruence with Previous Hummingbird Phylogenetic Studies

The phylogenetic estimate presented here is largely congruent with previous hummingbird phylogenetic studies. For example, the higher-level phylogenetic framework proposed by Bleiweiss et al. (1997) on the basis of a limited set of species and largely confirmed by Altshuler et al. (2004) is further corroborated here. Thus, our results are consistent with an informal higher-level hummingbird taxonomy that includes as major groupings the Hermits, Mangoes, Coquettes, Brilliants, Mountain Gems, Bees, and Emeralds (see Bleiweiss et al., 1997; Altshuler et al., 2004). Although this framework appears capable of accommodating most of hummingbird taxonomic diversity, there are at least four species that cannot be reconciled with these major groupings. These include the Crimson Topaz (*Topaza pella*), the White-necked Jacobin (*Florisuga mellivora*), the Black Jacobin (*Florisuga fusca*), and the Giant Hummingbird (*Patagona gigas*). Our phylogenetic analyses suggest basal phylogenetic positions for *Topaza* and *Florisuga* that throw into question the conventional taxonomic arrangement in which hummingbirds are divided into two putatively monophyletic subfamilies, Phaethornithinae (the hermits) and Trochilinae (all other hummingbirds). The early positions of *Topaza* and *Florisuga* in the linear sequence (e.g., Peters, 1945; Meyer de Schauensee, 1966) indicate that avian systematists have long suspected these taxa to be relatively basal and probably affiliated with Mangoes. Indeed, *Topaza* and *Florisuga mellivora* are known to have the type 2 condition of the tensor patagii brevis muscle (Zusi and Bentz, 1982), otherwise known only in members of the monophyletic Mangoes assemblage (Figs. 2, 3). Our results confirm that *Topaza* and *Florisuga* diverged early in hummingbird evolution but are surprising in that they suggest that they together represent the sister taxon of all other hummingbirds. However, we must emphasize that this basal phylogenetic arrangement is not strongly supported, and we cannot discount the possibility that the *Florisuga-Topaza* clade is the sister taxon of all remaining trochilines (all non-hermits) or even the sister taxon of the hermits. Although the basal phylogenetic position of *Florisuga* and *Topaza* is not strongly supported, the data do strongly reject their placement within either the Hermit or Mango clades, and we therefore conclude that this small clade must represent another primary branch within hummingbirds.

Our phylogenetic results indicate that the Giant Hummingbird, *Patagona gigas*, cannot be accommodated by the Bleiweiss et al. (1997) informal taxonomy. Our analyses suggest that *P. gigas* is the sister taxon of a large group comprised of the Emerald, Mountain Gem, and Bee clades. Thus, *P. gigas* appears to be a relatively old

lineage and a single, relatively unsuccessful (from the standpoint of species diversification) evolutionary experiment with gigantism among hummingbirds.

Historical Biogeography

Our data strongly support the hypothesis that hummingbirds originated in the South American lowlands, as suggested by Bleiweiss (1998). We base this conclusion on two primary findings. First, our ML ancestral state reconstruction analyses indicate that seven of the nine primary clades of hummingbirds, including at least the six most-basal clades, unequivocally originated in South America, with Mountain Gems and Bees representing the two possible exceptions (Fig. 5). Second, the common ancestors of the three most-basal hummingbird clades (Topazes, Hermits, and Mangoes) are unequivocally inferred to have had lowland distributions (Fig. 6). Although the interpretation of a lowland South American origin for hummingbirds and several of the more basal hummingbird clades is straightforward given that each node receives greater than 0.999 likelihood support for this reconstruction, there are interesting complications associated with clades nested more deeply in the tree. For example, our ML analyses strongly indicate that Mountain Gems originated in Central America (>0.99 likelihood support), with a few deeply nested species within this clade later expanding their ranges into North America (i.e., *Eugenes fulgens*) and South America (i.e., *Heliomaster longirostris*). The geographic origin of Bees, however, remains less certain. Bees can be divided into two primary clades, one of which primarily occurs in North and Central America, and the other primarily in South America (with a single Central American colonist). Bleiweiss (1998) argued that the ancestor of the South American Bee clade invaded from temperate Central/North America. Our analysis suggests a South American origin for this group, which is consistent with an invasion in the opposite direction (South to North America). However, the strength of support for our inference is sensitive to our rescaling of the hummingbird tree, with the proportion of the likelihood supporting a South American ancestor equally 0.781, 0.886, or 0.991 given mean tree lengths of 30.0, 13.0, or 1.0 (with 13.0 approximately representing the empirically estimated tree length based on our Bayesian phylogenetic analysis). Perhaps with more complete sampling of Bee hummingbirds, it will be possible to reconstruct the ancestral distribution of this group with greater certainty.

Although South America is the clear center of origin for the majority of hummingbird diversity, our phylogenetic hypothesis suggests approximately 30 to 34 independent dispersal events out of South America, 28 of which involve dispersal from South America into Central America. However, we note that 24 of these inferred dispersal events involve single species that are members of otherwise South American clades that have apparently extended their distributions into Central America, and occasionally into North America fairly recently. Indeed, only one subclade outside of Mountain Gems and Bees (a

clade of Central American Emeralds comprised of *Elvira*, *Eupherusa*, and *Microchera*) appears to have diversified to any large degree outside of South America. All other taxa with Central American, North American, or Caribbean distributions appear to be relatively recent arrivals (based on our current sampling). We have evidence for only one unambiguous dispersal event back into South America (from Central America: *Heliomaster longirostris*; but note that the ranges of the two *Heliomaster* species not considered in this study, *H. furcifer* and *H. squamosus*, are entirely within South America). These findings strongly suggest that assembly of the Central and North American hummingbird faunas has been strongly influenced by recent dispersal from South America, presumably since Panamanian uplift, and that the Central American hummingbird fauna, in particular, must have been much less diverse prior to this recent burst of dispersal into the region. In contrast, the South American fauna appears to have been assembled primarily from within.

We indicated above that the three most basal hummingbird clades had lowland origins. Indeed, among the principle clades of hummingbirds, only the Coquettes and Brilliants (which together comprise the Andean clade) appear to have had high-elevation common ancestors (Fig. 6). This finding is at odds with Bleiweiss (1998), who hypothesized that Coquettes and Brilliants each had lowland ancestors that invaded high elevation Andean habitats independently. However, we cannot discount Bleiweiss's (1998) hypothesis because phylogenetic uncertainty at the base of the Coquette clade leaves open the possibility that the Coquette common ancestor had a lowland distribution. Although Coquette monophyly is strongly supported in our analyses, the relative positions of the first two branches are weakly supported. Our ancestral state reconstructions are based on the Bayesian consensus topology, which has the high-elevation Andean genus *Heliangelus* placed as the sister taxon to the remaining members of the clade. However, if we switch the positions of *Heliangelus* and a lowland clade comprised of *Sephanoides*, *Discosura*, and *Lophornis*, the reconstruction is consistent with a low elevation, potentially non-Andean common ancestor of Coquettes (although the reconstructed ancestral minimum elevation for Coquettes would still be nearly 1300 m because of the topological proximity of Brilliants and the remaining high-elevation Coquettes). Because this critical basal node is weakly resolved, we cannot reconstruct the ancestral state for this clade with confidence.

We infer a lowland South American origin for hummingbirds, but our analyses of elevational occurrence show clearly that a substantial component of hummingbird diversity evolved in highland habitats (i.e., above 2000 m). However, most of the high elevation diversity is concentrated in the Andean clade (Coquettes + Brilliants). We performed separate squared-change parsimony reconstructions of minimum elevational occurrence (Fig. 6), mid-elevational occurrence, and maximum elevational occurrence, and all three analyses clearly support our contention that the three basal-most hummingbird clades (Topazes, Hermits, and Mangoes),
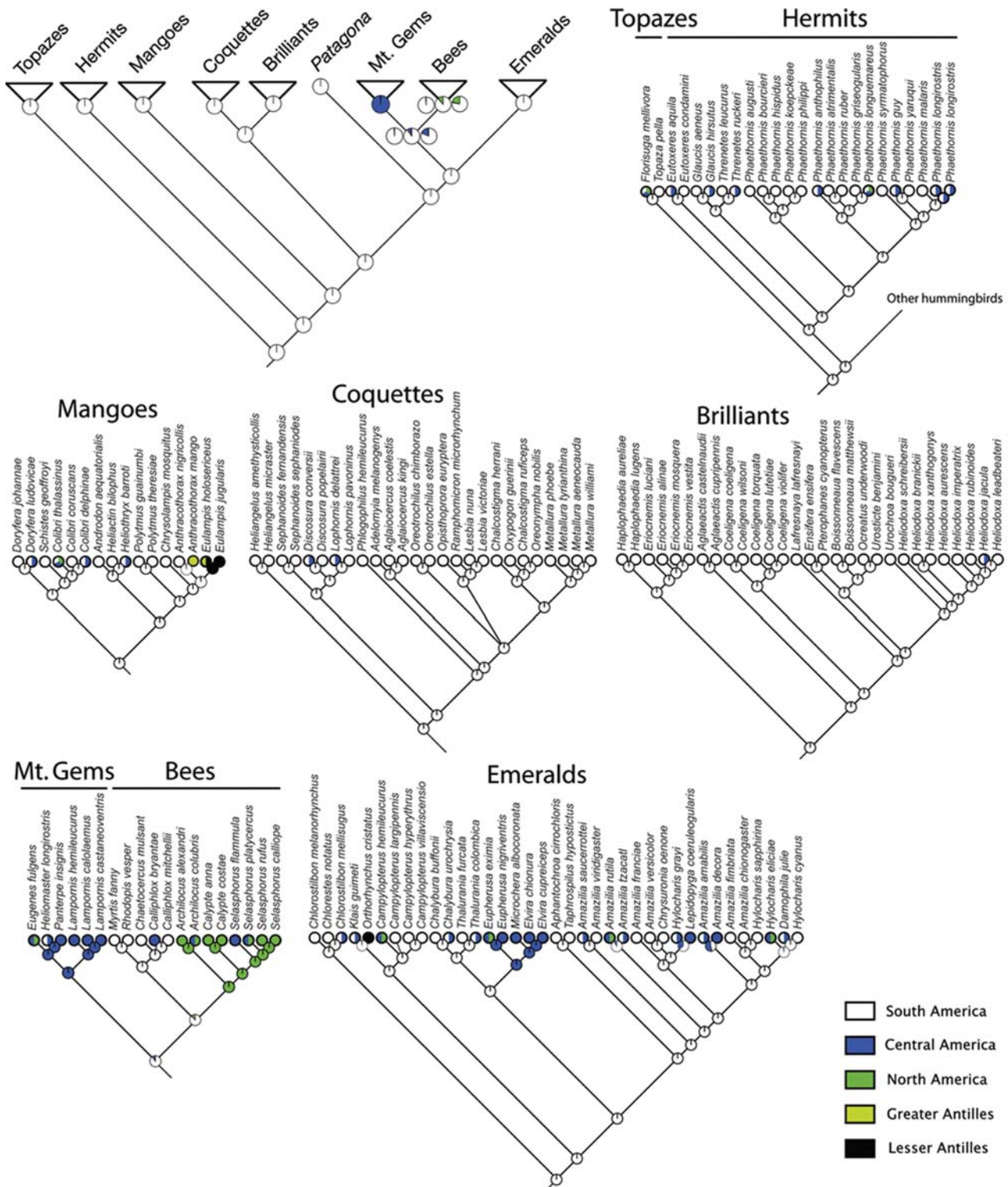
FIGURE 5. Maximum likelihood ancestral state reconstructions of primary landmass for each node on the 6B-partition Bayesian consensus tree. Pie diagrams at each node indicate the proportion of the maximum likelihood supporting alternative reconstructed character states. Reconstructions at nodes with one pie diagram were obtained in the analysis in which treelengths were drawn from a gamma distribution with a mean of 13.0 and variance of 5.0. The Bee and Mountain Gem + Bee nodes are represented by three diagrams because reconstructions at these two nodes were sensitive to tree length scaling. The left, center, and right diagrams summarize reconstructions based on analyses in which tree lengths were drawn from gamma distributions with means/variances of 1.0/0.5, 13.0/5.0, and 30.0/10.0.
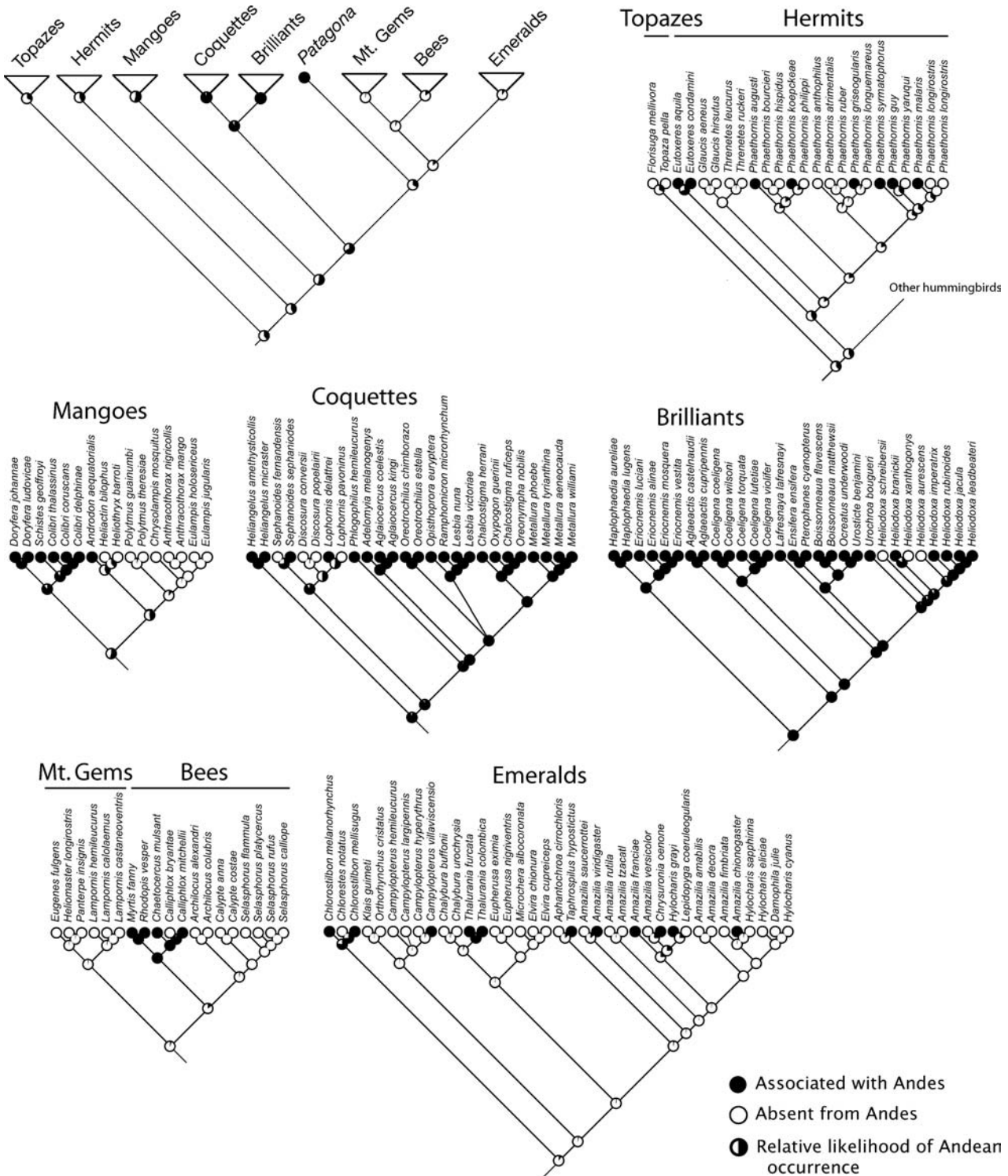
FIGURE 6. Maximum likelihood Markov model (Mk1) ancestral state reconstructions describing Andean versus non-Andean geographic associations. Pie diagrams indicate the relative likelihood for alternative reconstructions.

as well as the clade comprised of *Patagona*, Mountain Gems, Bees, and Emeralds had lowland common ancestors. Analyses of minimum elevational occurrence suggest that the common ancestors of these four clades had

minimum elevations in the lowlands (380 to 743 m, but we caution against the precise interpretation of elevation values, especially those below 1000 m), whereas the common ancestor of the Andean clade had a reconstructed
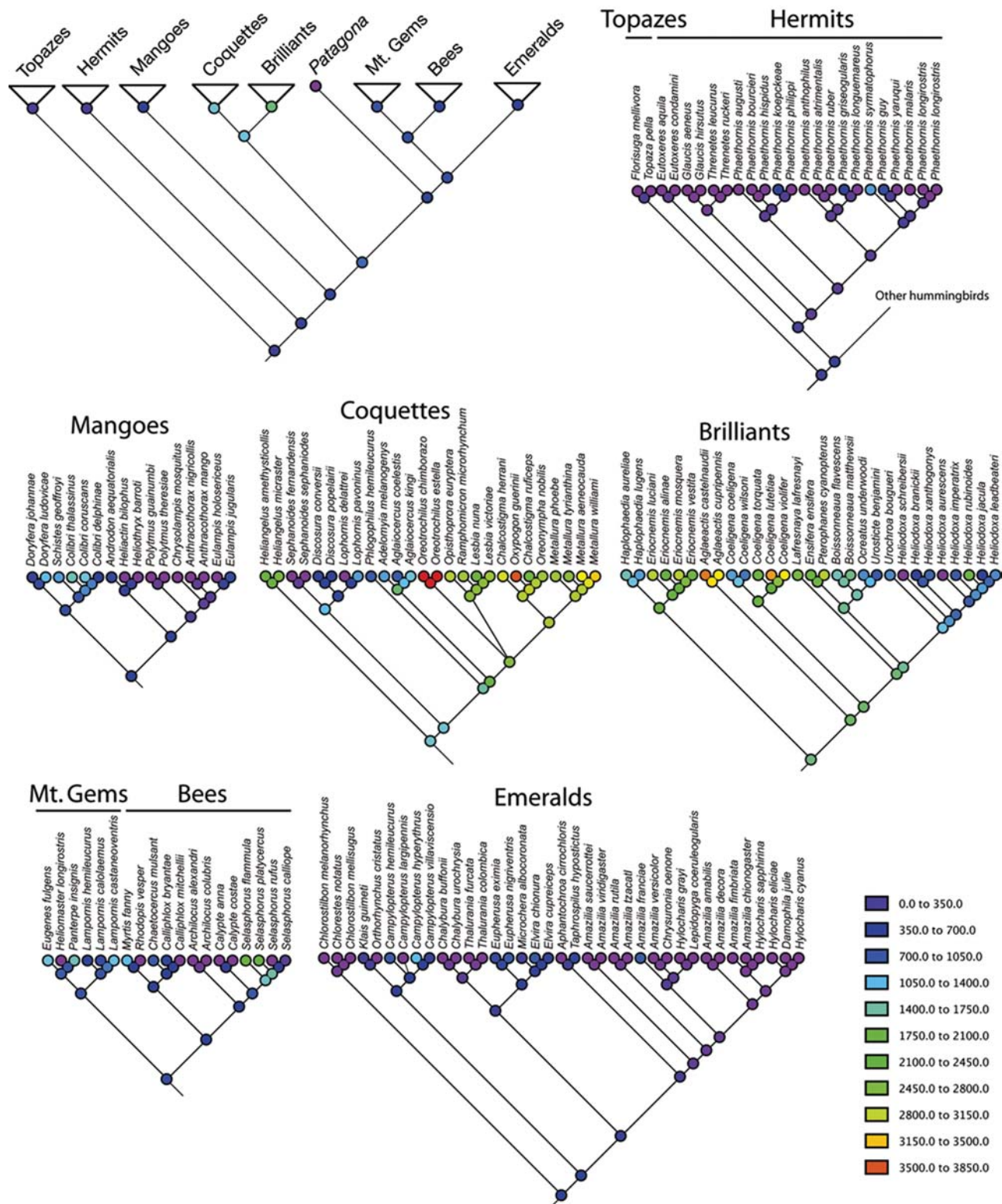
FIGURE 7. Weighted squared-change parsimony ancestral state reconstruction of minimum elevational occurrence for hummingbirds. Units are meters above sealevel.

minimum elevation of ~1234 m. This estimate is influenced by the lower elevations reconstructed for the non-Andean clades, and estimated elevations for many of the internal nodes within the Andean clade are even higher. The elevational trends within the Coquettes and Brilliants differ in that the more deeply nested branches in the Coquette clade tend toward higher elevational occurrences, whereas the more deeply nested nodes

within the Brilliants clade tend toward lower elevational occurrences.

Our investigation of hummingbird diversity in the Andes Mountains clearly indicates that hummingbird diversification has been intimately linked to this mountain range (Fig. 7). The Andean clade, in particular, includes a substantial percentage of extant hummingbird diversity and nearly every species in this appropriately named clade is associated with the Andes (48 of 55 Andean clade species included in this study have complete or partial Andean distributions). Furthermore, with the exception of Topazes and Mountain Gems, every major clade of hummingbirds has substantial Andean representation suggesting a complex pattern of dispersal and vicariance. Indeed, occurrence in the Andes is so prevalent that there is substantial likelihood support for an Andean common ancestor for all of the major clades except Mountain Gems, Bees, and Emeralds. Thus, although there is sufficient signal in the data to identify invasion of, or dispersal from, the Andes in specific instances (i.e., within Emeralds, Bees, Coquettes, and Brilliants), the reconstructions are essentially equivocal for the ancestor of hummingbirds and all of the basal-most nodes on the tree. Nevertheless, even though there are many nodes for which we cannot distinguish between Andean versus non-Andean ancestry, there is sufficient signal in the data to conservatively identify at least 10 invasions of, and five dispersal events from, these mountains, indicating the dynamic nature of hummingbird diversification in the Andes.

Finally, we note that a series of putative hummingbird fossils have been discovered in recent years in Oligicene deposits of Europe (Mayr, 2003, 2004, 2007). These discoveries were unexpected because extant hummingbirds are entirely restricted to the New World.

If the European fossils truly represent the earliest modern hummingbirds (Mayr, 2004), we might furthermore expect to find early diverging hummingbird lineages in North America as well. However, we have shown that extant hummingbirds are relatively recent colonizers of North America. Taken together, these observations suggest the intriguing possibilities that hummingbirds directly colonized South America from Europe (or elsewhere in the Old World) early in their history before going extinct in that region, or that early hummingbird lineages were once present in North America, disappeared, and later recolonized. Future divergence dating analyses based on more extensive sampling of extant hummingbirds should shed light on the temporal sequence of these fossil stem group trochilids and their colonization of South America.

REFERENCES

Abdo, Z., V. Minin, P. Joyce, and J. Sullivan. 2005. Accounting for uncertainty in the tree topology has little effect on the decision theoretic approach to model selection in phylogeny estimation. Mol. Biol. Evol. 22:691–703.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* Second international symposium on information theory (P. N. Petrov and F. Csaki, eds.). Akad. Kiado, Budapest.

Altekar, G., S. Dwarkadas, J. P. Huelsenbeck, and F. Ronquist. 2004. Parallel Metropolis-coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics 20:407–415.

Altshuler, D. L., and R. Dudley. 2002. The ecological and evolutionary interface of hummingbird flight physiology. J. Exp. Biol. 205:2325–2336.

Altshuler, D. L., R. Dudley, and J. A. McGuire. 2004. Resolution of a paradox: Hummingbird flight at high elevations does not come without a cost. Proc. Natl. Acad. Sci. USA 101:17731–736.

American Ornithologists' Union. 1998. Check-list of North American birds, 7th edition. American Ornithologists' Union, Washington, District of Columbia.

Anderson, F. E., and D. L. Swofford. 2004. Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. Mol. Phylogenet. Evol. 33:440–451.

Arevalo E., S.K. Davis, and J.W. Sites. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in Central Mexico. Syst. Biol. 43:387–418.

Bartholomew, G. A., and J. R. B. Lighton. 1986. Oxygen-consumption during hover-feeding in free-ranging Anna's Hummingbirds. J. Exp. Biol. 123:191–199.

Bleiweiss, R. 1998. Origin of hummingbird faunas. Biol J. Linn. Soc. 65:77–97.

Bleiweiss, R., J. A. W. Kirsch, and J. C. Matheus. 1997. DNA hybridization evidence for the principal lineages of hummingbirds (Aves: Trochilidae). Mol. Biol. Evol. 14:325–343.

Brandley, M. C., A. Schmitz, and T. W. Reeder. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst. Biol. 54:373–390.

Buschbom, J., and D. Barker. 2006. Evolutionary history of vegetative reproduction in *Porpidia s.l.* (lichen-forming Ascomycota). Syst. Biol. 55:471–484.

Castoe, T. A., and C. L. Parkinson. 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). Mol. Phylogenet. Evol. 39:91–110.

Castoe, T. A., M. M. Sasa, and C. L. Parkinson. 2005. Modeling nucleotide evolution at the mesoscale: The phylogeny of the Neotropical pitvipers of the *Porthidium* group (Viperidae: Crotalidae). Mol. Phylogenet. Evol. 37:881–898.

Chai, P., and R. Dudley. 1995. Limits to vertebrate locomotor energetics suggested by hummingbirds hovering in heliox. Nature 377:722–725.

Chai, P., and R. Dudley. 1996. Limits to flight energetics of hummingbirds hovering in hypodense and hypoxic gas mixtures. J. Exp. Biol. 199:2285–2295.

Chang, J. T. 1996. Full reconstruction of Markov models on evolutionary trees: Identifyability and consistency. Math. Biosci. 137:51–73.

Chapman, F. M. 1925. Remarks on the life zones of northeastern Venezuela with descriptions of new species of birds. Am. Mus. Novit. 191:1–15.

Chapman, F. M. 1926. The distribution of bird-life in Ecuador. Bull. Am. Mus. Nat. Hist. 55:1–784.

Cracraft, J. 1985. Historical biogeography and patterns of differentiation within the South American avifauna: Areas of endemism. Pages 49–84 in Neotropical ornithology (P. A. Buckley, M. S. Foster, E. S. Morton, R. S. Ridgely, and F. G. Buckley, eds.). Ornith. Mono. No. 36. American Ornithologists' Union, Washington, District of Columbia.

Cracraft, J., F. K. Barker, M. Braun, J. Harshman, G. J. Dyke, J. Feinstein, S. Stanley, A. Cibois, P. Schikler, P. Beresford, J. García-Moreno, M. D. Sorenson, T. Yuri, and D. P. Mindell. 2004. Phylogenetic relationships among modern birds (Neornithes): Toward an avian tree of life. Pages 468–489 in Assembling the Tree of Life (J. Cracraft and M. J. Donoghue, eds.). Oxford University Press, Oxford, UK.

Dickinson, E. 2003. The Howard & Moore complete checklist of the birds of the world, 3rd edition. Princeton University Press, Princeton, New Jersey.

Dziedzioch, C., A.-D. Stevens, and G. Gottsberger. 2003 The hummingbird plant community of a tropical montane rain forest in southern Ecuador. Plant Biol. 5:331–337.

Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Felsenstein, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

Gaut, B. S., and P. O. Lewis. 1995. Success of maximum likelihood phylogeny inference in the four-taxon case. Mol. Biol. Evol. 12:152–162.

Grant, K. A., and V. Grant. 1968. Hummingbirds and their flowers. Columbia University Press, New York.

Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. Syst. Biol. 44:17–48.

Huelsenbeck, J. P., and D. M. Hillis. 1993. Success of phylogenetic methods in the four-taxon case. Syst. Biol. 42:247–264.

Huelsenbeck, J. P., and B. Rannala. 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53:904–913.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

Kass, R. E., and A. E. Raftery. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773–795.

Krömer, T., and M. Kessler. 2006. Distribution and flowering ecology of bromeliads along two climatically contrasting elevational transects in the Bolivian Andes. Biotropica 38:183–195.

Kumazawa, Y., and M. Nishida. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. 37:380–398.

Lack, D. 1973. The number of species of hummingbirds in the West Indies. Evolution 27:326–327.

Lack, D. 1976. Island biology illustrated by the land birds of Jamaica. University of California Press, Berkeley, California.

Lartillot, N., and H. Philippe. 2006. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55:195–207.

Lemmon, A. E., and E. C. Moriarty. 2004. The importance of proper model assumption in Bayesian phylogenetics. Syst. Biol. 53:265–277.

Lutzoni, F. M., M. Pagel, and V. Reeb. 2001. Major fungal lineages are derived from lichen symbiotic ancestors. Nature 411:937–940.

Macey, J. R., and A. Verma. 1997. Homology in phylogenetic analysis: Alignment of transfer RNA genes and the phylogenetic position of snakes. Mol. Phylogenet. Evol. 7:272–279.

Maddison, D. R., and W. P. Maddison. 2003. MacClade, version 4.06. Sinauer Associates, Sunderland, Massachusetts.

Maddison, W. P. 1991. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. Syst. Zool. 40:304–314.

Maddison, W. P., and D. R. Maddison. 2005. Mesquite: A modular system for evolutionary analysis, version 1.06. http://www.mesquiteproject.org.

Marshall, D. C., C. Simon, and T. R. Buckley. 2006. Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. Syst. Biol. 55:993–1003.

Mayr, G. 2003. Phylogeny of early Tertiary swifts and hummingbirds (Aves: Apodiformes). Auk 120:145–151.

Mayr, G. 2004. Old World fossil record of modern-type hummingbirds. Science 304:861–864.

Mayr, G. 2007. New specimens of the early Oligocene Old World hummingbird Eurotrochilus inexpectatus. J. Ornithol. 148:105–111.

Meyer de Schauensee, R. 1966. The species of birds of South America and their distribution. Livingston Publishing Co., Narbeth, Pennsylvania.

Minin, V, Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogeny estimation. Syst. Biol. 52:674–683.

Moyle, R. G., J. Cracraft, M. Lakim, J. Nais, and F. H. Sheldon. Reconsideration of the phylogenetic relationships of the enigmatic Bornean Bristlehead (Pityriasis gymnocephala). Mol. Phylogenet. Evol. 39:893–898.

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.

Pagel, M., and A. Meade. 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene-sequence or character-state data. Syst. Biol. 53:571–581.

Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. Syst. Biol. 53:571–581.

Parker, T. A., III, T. S. Schulenberg, G. R. Graves, and M. J. Braun. 1985. The avifauna of the Huancabamba region, northern Peru. Pages 169–197 in Neotropical ornithology (P. A. Buckley, M. S. Foster, E. S. Morton, R. S. Ridgely, and F. G. Buckley, eds.) Ornithol. Monogr. 36. American Ornithologists' Union, Washington, D.C.

Parker, T. A., III, D. F. Stotz, and J. W. Fitzpatrick. 1996. Ecological and distributional databases. Pages 113–460 in Neotropical birds, ecology and conservation (D. F. Stotz, J. W. Fitzpatrick, T. A. Parker, III, and D. K. Moskovits, eds.). University of Chicago Press, Chicago, Illinois.

Peters, J. L. 1945. Check-list of the birds of the world, volume 5. Harvard University Press, Cambridge, Massachusetts.

Posada, D., and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53:793–808.

Posada, D., and K. A. Crandall. 1998. ModelTest: Testing the model of DNA substitution. Bioinformatics 14:817–818.

Prychitko, T. M., and W. S. Moore. 1997. The utility of DNA sequences of an intron from the β-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). Mol. Phylogenet. Evol. 8:193–204.

Rahbek, C., and G. R. Graves. 2000. Detection of macro-ecological patterns in South American hummingbirds is affected spatial scale. Proc. R. Soc. Lond. B 267:2259–2265.

Rannala, B. 2002. Identifiability of parameters in MCMC Bayesian inference of phylogeny. Syst. Biol. 51:754–760.

Robbins, M.B., R. S. Ridgely, T. S. Schulenberg, F. B. Gill. 1987. The avifauna of the Cordillera-De-Cutucu, Ecuador, with comparisons to other Andean localities. Proc. Acad. Nat. Sci. Phila. 139:243–259.

Ronquist, F. 1996. DIVA, version 1.1. Computer program and manual available by anonymous FTP from Uppsala University (ftp.uu.se or ftp.systbot.uu.se).

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Schluter, D., T. Price, A. Ø. Mooers, and D. Ludwig. 1997. Likelihood of ancestral states in adaptive radiation. Evolution 51:1699–1711.

Schuchmann, K. L. 1999. Family Trochilidae (hummingbirds). Pages 468–680 in Handbook of the birds of the world, volume 5, Barn-owls to hummingbirds (J. del Hoyo, A. Elliot, and J. Sargatal, eds.). Lynx Edicions, Barcelona.

Schwarz, G. 1978. Estimating the dimensions of a model. Ann. Stat. 6:461–464.

Shapiro L. H., and J. P. Dumbacher. 2001. Adenylate kinase intron 5: A new nuclear locus for avian systematics. Auk 118:248–255.

Smith, S. A., P. R. Stephens, and J. J. Wiens. 2005. Replicate patterns of species richness, historical biogeography, and phylogeny in Holarctic treefrogs. Evolution 59:2433–2450.

Sorenson, M. D., J. C. Ast, D. E. Dimcheff, T. Yuri, and D. P. Mindell. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. Mol. Phylogenet. Evol. 12:105–114.

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

Steel, M. 2005. Should phylogenetic models be trying to "fit an elephant"? Trends Genet. 21:307–309.

Stiles, F. G. 1981. Geographical aspects of bird-flower coevolution, with special reference to Central America. Ann. Missouri Bot. Garden 68:323–351.

Stiles, F. G., and D. J. Levy. 1994. Birds of La Selva and vicinity. Pages 384–393 in La Selva: Ecology and natural history of a Neotropical rain forest (L. A. McDade, K. S. Bawa, H. A. Hespenheide, and G. S. Hartshorn, eds.). University of Chicago Press, Chicago, Illinois.

Suarez, R. K. 1998. Oxygen and the upper limits to animal design and performance. J. Exp. Biol. 201:1065–1072.

Suarez, R. K., J. R. B. Lighton, G. S. Brown, and O. Mathieu-Costello. 1991. Mitochondrial respiration in hummingbird flight muscles. Proc. Natl. Acad. Sci. USA 88:4870–4873.

Sullivan, J., and P. Joyce. 2005. Model selection in phylogenetics. Annu. Rev. Ecol. Evol. Syst. 36:445–466.

Sullivan, J., and D. L. Swofford. 1997. Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. J. Mammal. Evol. 4:77–86.

Sullivan, J., and D. L. Swofford. 2001. Should we use model-based methods of phylogenetic inference when we know assumptions about among-site rate variation and nucleotide substitution pattern are violated? Syst. Biol. 50:723–729.

Swofford, D. L. 2003. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Swofford, D. L., P. J. Waddell, J. P. Huelsenbeck, P. G. Foster, P. O. Lewis, and J. S. Rogers. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. Syst. Biol. 50:525–539.

Temeles E. J., and W. J. Kress. 2003. Adaptation in a plant-hummingbird association. Science 300:630–633.

Wiens, J. J., J. W. Fetzner, Jr., C. L. Parkinson, and T. W. Reeder. 2005. Hylid frog phylogeny and sampling strategies for speciose clades. Syst. Biol. 54:719–748.

Wilgenbusch J. C., D. L. Warren, and D. L. Swofford. 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. http://ceb.csit.fsu.edu/awty.

Zusi, R. L., and G. D. Bentz. 1982. Variation in a muscle in hummingbirds and swifts and its systematic implications. Proc. Biol. Soc. Wash. 95:412–420.

Andean Hillstar (*Oreotrochilus estella stolzmanni*) female at 3600 m elevation, Lambayeque, Peru, photographed by Andrew B. Johnson. Members of this genus are physiological marvels known to occur up to 5000 m. The genus is placed in a clade within the coquette group in which all the members are restricted to middle or high elevations in the Andes.