

## Species Delimitation Using a Combined Coalescent and Information-Theoretic Approach: An Example from North American *Myotis* Bats

BRYAN C. CARSTENS<sup>1,\*</sup> AND TANYA A. DEWEY<sup>2</sup>

<sup>1</sup>Department of Biological Sciences, Louisiana State University, 202 Life Sciences Building, Baton Rouge, LA 70808, USA; and <sup>2</sup>Department of Ecology and Evolutionary Biology, Museum of Zoology, University of Michigan, 1109 Geddes Avenue, Ann Arbor, MI 48109-1079, USA;

\*Correspondence to be sent to: Department of Biological Sciences, Louisiana State University, 202 Life Sciences Building, Baton Rouge, LA 70808, USA; E-mail: carstens@lsu.edu.

Received 20 May 2009; reviews returned 18 August 2009; accepted 11 December 2009

Associate Editor: Marshal Hedin

**Abstract.**—Coalescent model-based methods for phylogeny estimation force systematists to confront issues related to the identification of species boundaries. Unlike conventional phylogenetic analysis, where species membership can be assessed qualitatively after the phylogeny is estimated, the phylogenies that are estimated under a coalescent model treat aggregates of individuals as the operational taxonomic units and thus require a priori definition of these sets because the models assume that the alleles in a given lineage are sampled from a single panmictic population. Fortunately, the use of coalescent model-based approaches allows systematists to conduct probabilistic tests of species limits by calculating the probability of competing models of lineage composition. Here, we conduct the first exploration of the issues related to applying such tests to a complex empirical system. Sequence data from multiple loci were used to assess species limits and phylogeny in a clade of North American *Myotis* bats. After estimating gene trees at each locus, the likelihood of models representing all hierarchical permutations of lineage composition was calculated and Akaike information criterion scores were computed. Metrics borrowed from information theory suggest that there is strong support for several models that include multiple evolutionary lineages within the currently described species *Myotis lucifugus* and *M. evotis*. Although these results are preliminary, they illustrate the practical importance of coupled species delimitation and phylogeny estimation. [Coalescent theory; species delimitation; species tree phylogeny estimation.]

One of the foremost goals of evolutionary biology is to understand the processes that promote speciation. In order to identify these processes, biologists must first recognize and delimit nascent evolutionary lineages (Sites and Marshall 2004; Wiens 2007). Genetic data at the interface between populations and species are generally useful for delimitation efforts, particularly gene trees from neutral loci (Harrison 1998; Templeton 2001). Because the pattern of allele coalescence is stochastic and can be defined in a probabilistic manner (Tajima 1983; Takahata and Nei 1985; Hudson et al. 1987), the rate that ancestral polymorphism is lost in a given lineage provides valuable information concerning both temporal divergence between sister lineages and the branching pattern of the phylogeny. Gene trees can be employed in delimitation efforts in 2 ways. For single-locus data, qualitative assessments of lineage boundaries can be derived from the genealogy estimate (Pellegrino et al. 2005). However, because any particular gene tree may not reflect the actual pattern of lineage splitting and divergence, this approach can be misled by coalescent stochasticity (Hudson and Coyne 2002). Data from multiple unlinked loci can mitigate the effects of the randomness inherent to genetic drift, and the acquisition of these data has led to a dramatic shift in how the phylogenies of recently diverged clades are estimated.

A number of factors can produce incongruence between species trees and gene trees (Maddison 1997), and indeed there are combinations of species tree branch lengths where the most common gene tree is expected to be incongruent with the species tree (Degnan and

Rosenberg 2006; Rosenberg and Tao 2008). These theoretical and practical difficulties with concatenation have spurred the incorporation of coalescent models into phylogenetic methods, a development that has been described as one of the most exciting recent developments in systematics (Page and Sullivan 2008; Edwards 2009).

Phylogenetic inference near the species level necessarily incorporates aspects of population genetic theory and methodology because the genetic forces acting within a population, such as genetic drift, selection, and migration, may each play an important role in speciation (Maddison and Knowles 2006). Several recently introduced phylogenetic methods operate under the assumption that genetic drift has produced the incongruence between gene trees and species trees (Maddison W.P. and Maddison D.R. 2004; Degnan and Salter 2005; Ané et al. 2007; Edwards et al. 2007; Liu and Pearl 2007; Oliver 2008; Kubatko et al. 2009). Relative to the delimitation of evolutionary lineages, the most important difference between these approaches and conventional methods of phylogenetic inference is the shift in the operational taxonomic units (OTUs). Rather than using a single or several representative individuals as exemplars, the species or population lineages are used as the OTUs and multiple alleles are sampled within each lineage. Thus, in addition to incorporating coalescent models of population processes in phylogenetic inference, methods of species tree estimation also model the membership of individuals to evolutionary lineages. This novel approach enables the relationships among these lineages to be inferred directly and also allows the membership of individual samples in these lineages to

be explored directly. Knowles and Carstens (2007) outlined the first approach to conduct lineage delimitation using a method for species tree inference:

For a given pair of terminal sister lineages two species trees are compared. In the first phylogeny, the sister lineages  $[A, B]$  are separated by a branch of some length  $X$ , in the alternate phylogeny the branch that separates these lineages is collapsed (e.g., separated by a branch with a length of zero). The probability of the gene trees given the two species trees is calculated, and significance is assessed with a likelihood ratio test. If the null is rejected, then the model where the sister lineages are separated by a branch of some length better fits the data, suggesting that the lineages are independent.

The Knowles–Carstens approach to delimitation represents a shift in how genetic data can be used to delimit lineages (de Querioz 2007). Rather than equating gene trees with a species tree or basing species status on some genetic threshold, the relationship between the gene trees and the lineage history is modeled probabilistically using coalescent theory (Hudson 1991). Adopting this explicit model-based approach also avoids problems with species delimitation that result when genetic thresholds are applied to genetic data—the detection biases arising from the timing and method of speciation and failure of any threshold to take into account the stochastic variance associated with genetic processes. However, the Knowles–Carstens approach has practical shortcomings, some of which are related to its use of the program COAL in an application not intended by Degnan and Salter (2005). Since its publication, additional methods for estimating species phylogeny have been developed, and at least 2 of these (Bayesian estimation of species trees [BEST], species tree estimation using maximum likelihood [STEM]) exhibit considerable potential for the delimitation of evolutionary lineages. Here, we use these programs to explore lineage boundaries in 10 described subspecies from the *Myotis lucifugus*/western long-eared *Myotis* clade.

#### Background Information North American *Myotis*

Molecular sequence data from mitochondrial genes are inconsistent with described species boundaries in the *M. lucifugus*/western long-eared *Myotis* clade of North American bats. Although this clade is consistently monophyletic in broadscale analyses of *Myotis* phylogeny (Ruedi and Mayer 2001; Stadelmann et al. 2007), relationships within it are difficult to infer. For example, the genealogy from cytochrome *b* suggests that *M. lucifugus*, *M. thysanodes*, and *M. evotis* are each paraphyletic (Dewey 2006). Results of phylogeographic analyses within this lineage are consistent with similar investigations from other animals, which demonstrate that there can be dramatic incon-

gruence between recognized species boundaries and clades contained within the genealogies of particular genetic loci (Riddle et al. 2000; Shaw 2000; Funk and Omland 2003). When confronted with paraphyly at mitochondrial DNA loci, researchers have been conditioned to gather sequence data from additional nuclear loci (Brumfield et al. 2003), but to date it is unclear how researchers can best use these data to assess species limits.

Taxonomy within North American *Myotis* has been influenced by the idea that morphological similarity is the result of close relationship, with species descriptions often based on regional differences in fur and membrane color and relative sizes of body measurements. The western long-eared *Myotis* group has traditionally been recognized as including *M. evotis*, *M. keenii*, *M. milleri*, and *M. thysanodes*, as well as the southwestern species *M. auriculus*, and a largely eastern species *M. septentrionalis* (Baker and Stains 1955; Koopman 1963). However, molecular phylogenies demonstrate that the western long-eared *Myotis* group is closely related to *M. lucifugus* and that *M. auriculus* and *M. septentrionalis* are distantly related (Dewey 2006; Stadelmann et al. 2007). All available evidence suggest that *M. lucifugus* and the western long-eared *Myotis* species form a monophyletic clade that is well supported (e.g., bootstrap values of 100 and Bayesian posterior probabilities of 1.0; Stadelmann et al. 2007), but relationships within this clade remained unresolved.

At this stage of our research, conducting thorough phylogeographic investigations into this widespread and complex clade requires some understanding of species boundaries. Our aim for this investigation is 1) to determine if described species are monophyletic and 2) to understand how western long-eared *Myotis* are related to *M. lucifugus*. Accomplishing this goal requires that we explore the utility of probabilistic approaches to species delimitation as well as other assumptions that are common to the model-based approaches to species phylogeny estimation that we utilize here, particularly that the genealogies are drawn from neutral loci without any internal recombination and evolve in a clocklike manner.

#### METHODS

##### Collection of Samples

Tissue specimens were obtained through field sampling by Dewey and other field workers for over 200 individuals (see Dewey 2006). Animals were identified to species in the field and confirmed at the University of Michigan Museum of Zoology. We selected 34 individuals from among the vouchers provided by Dewey (Table 1). Samples were chosen from regions where assignment to described subspecies was likely to be unambiguous (Supplementary Table 1, available from <http://www.sysbio.oxfordjournals.org>). On the basis of results from Stadelmann et al. (2007), *M. volans* was chosen as an outgroup.

TABLE 1. Geographic distributions of *M. lucifugus*/western long-eared *Myotis* subspecies

Lineage	Distribution	No.
<i>M. evotis chrysonotus</i>	Interior BC, SK, AB, MT, W ND, W SD, WY, CO, UT, NV, CA east of the Sierras, SE OR, S ID	6
<i>M. e. jonesorum</i>	NE AZ, NW NM	2
<i>M. keenii</i>	S AK, BC, coastal WA	3
<i>M. thysanodes thysanodes</i>	Central N Mexico through Sierra Madres to NM, AZ, CO, UT, ID, NV, E & S CA	2
<i>M. t. pahasapensis</i>	E WY, SW SD, W NE	1
<i>M. t. vespertinus</i>	Coastal N CA, OR, WA, BC	1
<i>M. lucifungus lucifugus</i>	E North America (NF/NS south to GA, AL, MS), west to OK, E KS, NE, SD, ND, MB, AB, SK, YT, AK	7
<i>M. l. alascensis</i>	Coastal CA, OR, WA, BC	4
<i>M. l. carissima</i>	W SD, W ND, MT, ID, E WA, E OR, NV, UT, CO, E CA, N AZ, N NM	3
<i>M. l. relictus</i>	Central Sierras, W NV, ID, SW MT	4
<i>M. volans</i>	Baja Cal, N. Mexico, AZ, NM, CA, NV, OR, WA, ID, MT, SD, ND, AB, BC to SE AK,	1

Note: States and provinces are abbreviated with their postal abbreviations.

### Collection of Molecular Data

Mitochondrial sequence data were generated from the cytochrome *b* gene by Dewey (2006). We developed nuclear markers as follows: First, we downloaded sequence data from 6 *M. lucifugus* BAC libraries collected by the Large-Scale Genome Sequencing Program of the ENCODE consortium. Sequence data were generated by 1) downloading genomic data from regions ENm003, ENr322, ENr221, and ENm004, 2) identifying open reading frames (ORFs) in each, 3) picking pairs of ORFs that flanked regions consisting of 500–800 bp of non-ORFs, 4) removing loci that exhibited evidence of short interspersed nuclear elements (SINEs) (Ray et al. 2007) or heliobats (Pritham and Freschotte 2007), 5) using PRIMER<sup>3</sup> (Rozen and Skaletsky 2000) to design primers that were located in the flanking ORF regions and amplified across the putative non-ORF (e.g., similar to an exon primed intron-crossing [EPIC] approach) (Palumbi and Baker 1994; Creer 2007), and 6) screening primers using a single individual each of *M. lucifungus lucifugus*, *M. evotis*, and *M. thysanodes*. Introns were targeted because we hoped to identify loci with a sufficient number of variable sites to estimate gene trees with high confidence. Thirty-six primer pairs were tried, 17 amplified across the screening set. Of these 14 sequenced well and 12 appeared to be single copy after polymerase chain reaction (PCR) subcloning. Sequence data were generated from 6 of these loci (selected at random) from 32 other samples.

Sequence data were edited and aligned using Sequencer 4.8 (GeneCodes, Ann Arbor, MI). All individuals were sequenced in both directions. Phase of alleles was determined in 2 ways. For some combinations of samples and loci that did not produce reads with high-quality scores (>95%) when sequenced directly, we used PCR subcloning with proofreading *Taq* and established phase by sequencing 8 clones per sample. In this manner, over one-third of the sequences were generated by subcloning. For samples and loci that we were able to sequence directly with high-quality scores, we used Phase 2.1 (Stephens et al. 2001) after including alleles identified via subcloning. Alignment was conducted using Clustal 2.1 (Larkin et al. 2007) and checked and adjusted manually using MacClade

4.05 (Maddison W.P. and Maddison D.R. 2004). We searched for evidence of recombination within each locus using HyPhy (Kosakovsky-Pond et al. 2005, 2006) and TOPALi (Milne et al. 2004). We tested for selection using the Hudson–Kreitman–Aguade (HKA) test (Hudson et al. 1987) implemented in DnaSP (Rozas J. and Rozas R. 1999).

### Estimation of Gene Trees

MacClade was used to identify redundant alleles for each locus, and we removed these alleles for subsequent estimation of gene trees. Models of sequence evolution were identified using DT-Modsel (Minin et al. 2003), and gene trees were estimated using maximum likelihood (ML) in PAUP\* (Swofford 2002). Nodal support was assessed using ML and 100 bootstrap replicates with maxtrees = 1, and a likelihood ratio test (LRT) was used to test the molecular clock hypothesis.

### Species Tree Estimation

Three approaches were used to estimate the species tree from gene trees. First, the topology was inferred by minimizing the deep coalescence (MDC) using Mesquite. For this analysis, subtree pruning and regrafting (SPR) branch swapping and maxtrees = 1000 were used as settings in Mesquite. To generate nodal support, we used the AUGIST module (Oliver 2008) to estimate the species tree using topologies saved during nonparametric bootstraps (above).

We also used 2 model-based approaches to estimate the species tree. Recently, Kubatko et al. (2009) introduced an approach for STEM that estimates the species tree from samples of observed gene trees, under the assumption that discordance is produced by the coalescent process alone. STEM can produce an analytically derived ML species tree with branch lengths that is similar to the GLASS tree (Mossel and Vigoda 2005) when  $\theta$  is equal across branches and is a consistent estimator of the species tree when the gene trees are estimated without error (Kubatko et al. 2009). Thus, if the ML estimate of the species tree is the primary goal of an analysis, then the only uncertainty that must be considered by the researcher is the phylogenetic error associated with gene tree estimation. STEM assumes that the



gene trees are sampled from loci that are evolving in a clocklike manner because the analytical calculations in STEM are based on estimation of species divergence time. Consequently, we conducted STEM analyses using 1) only those loci that did not violate the molecular clock and 2) all loci, with branch length optimization constrained to be clocklike, which may be a reasonable option for non-clocklike data. Sequence data from cytochrome *b* was scaled in STEM analyses to reflect the difference in  $N_e$  between mitochondrial and autosomal loci. We omitted data from locus 681a in the STEM analyses because we were unable to amplify this locus in *M. volans*.

Species tree estimation in STEM assumes that  $\theta = 4N_e\mu$  is constant across lineages, and this value is provided by the user. Obtaining this estimate is potentially complicated by the uncertainty in species boundaries, for example, we might expect a larger value of  $\theta$  when it is estimated from species as opposed to subspecies because larger aggregates of samples generally contain more genetic diversity. Using Migrate-n (Beerli and Felsenstein 2001), we estimated  $\theta$  for all species, as well for selected subspecies. Migrate-n analyses followed search strategies described previously (Knowles et al. 2007).

BEST (Liu and Pearl 2007; Edwards et al. 2007) was also used to estimate the posterior distribution of species tree space. BEST is a Bayesian approach that approximates the posterior distribution of species trees from the posterior distributions of gene trees and therefore estimates branch lengths of both the species trees and the gene trees. To date, BEST has been applied to a few empirical data sets, but early results (Belfiore et al. 2008; Brumfield et al. 2008; Linnen and Farrell 2008) and performance evaluations (Edwards et al. 2007; Liu et al. 2008) are promising. BEST analyses were conducted using duplicate runs with 101,000,000 generations, a sample frequency of 1000 generations, and 2 coupled Markov chains. Prior settings included an inverse gamma (4.0, 0.001) and a uniform mutation prior (0.5, 2.0) with site-specific models of sequence evolution. Due to missing data in locus 681a, we ran all BEST runs with only 6 loci. Multiple BEST runs were conducted, including pairwise combinations of runs that treated species and subspecies as OTUs and those using clocklike loci and all loci.

We also concatenated the data across loci and estimated the phylogeny using MrBayes 3.0 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). For this analysis, alleles were chosen at random from each individual and locus and the analysis was conducted once with each set of alleles using a partitioned model that matched those selected above for each locus. For reasons that include both phylogenetic and coalescent error, concatenation across loci is not expected to be a reliable method for inferring the species phylogeny (Degnan and Rosenberg 2006; Kolaczkowski and Thornton 2006; Carstens and Knowles 2007; Kubatko and Degnan 2009).

### Tests for Lineage Delimitation

We sampled individuals from 10 described subspecies. The likelihood of all hierarchical permutations of these lineages was calculated such that we evaluated every combination of subspecies within any species and compared these combinations with those within the other species. There are 150 such combinations: the product of 15 permutations of the 4 subspecies within *M. lucifugus* (e.g., all subspecies treated as separate lineages, all subspecies combined into one lineage, 6 combinations of 3 lineages, and 7 combinations of 2 lineages), 5 permutations of the 3 subspecies within *M. thysanodes*, and 2 permutations of the 2 subspecies within *M. evotis*. Note that although we limited the permutations of lineage membership to those within described species, we did nothing to constrain the species phylogeny (e.g., paraphyletic species were allowed). Likelihoods were evaluated using LRTs following Knowles and Carstens (2007) for each of the above subspecies. Because this test uses estimates of topology and branch lengths of the species tree but assumes that estimated gene trees reflect the actual history of allelic coalescence, we also explored the degree to which uncertainty in the gene tree estimates complicates delimitation efforts. To accomplish this, we assessed lineage distinctiveness using BEST by computing the Bayes factor (Kass and Raftery 1995) between the posterior distributions of a BEST analysis where 2 lineages are separated and one where they are collapsed. This approach considers both the topology and the branch lengths of species trees, the properties most relevant to the general hypothesis outlined above, and accounts for uncertainty inherent to the estimation of gene trees. It is also conducted in a statistical framework that determines whether species trees with separate or collapsed lineages best fit the data. However, computing Bayes factors for all above comparisons would require 151 BEST runs, a computational commitment beyond our resources. As a result, these tests were conducted on selected lineages.

In a complex system such as the *M. lucifugus*/western long-eared *Myotis* clade, species delimitation cannot proceed using only LRTs because results may be difficult to interpret. As is the case in most statistical hypothesis testing approaches, the *P* value represents the probability that the test statistic (here twice the difference in likelihood scores between these models) is greater than expected given that the null hypothesis is true. Conflicting models that each fail to reject the null are impossible to rank. Additionally, the likelihoods of a particular grouping within one species will be influenced by the assumed lineage composition of another species. For these reasons, we also calculated the Akaike information criterion (AIC) scores (Akaike 1973) and model likelihoods (Burnham and Anderson 1998) of all models and evaluated these probabilities following the conceptual logic of information theory (Kullback 1959; Anderson 2008). In essence, we are evaluating multiple simultaneously.

*Effects of Incorrect Assignment of Individuals to Lineages on Species Tree Estimation*

Although STEM and BEST are new approaches to estimating phylogenies, they are firmly rooted in the tradition of model-based methods used by geneticists to correct for multiple nucleotide substitutions (Jukes and Cantor 1969). As with earlier approaches, the assumptions of these models are likely to be critical to the performance of these methods. For example, certain types of gene flow can degrade the accuracy of species tree estimates (Eckert and Carstens 2008), and species tree estimates improve when more individuals are sampled within lineages (McCormack et al. 2009). One assumption of species tree methods that has not been explored is the effect of incorrect assignment of samples to the aggregates that serve as the OTUs; this could lead to inaccurate estimation of the topology, species divergence times, or both. We conducted a small simulation analysis to explore this question. Genealogies were simulated with *ms* (Hudson 2002) under 2 species trees. In the first, 5 lineages were separated by branch lengths sufficient to allow for sorting of ancestral polymorphism (Fig. 1a). The second phylogeny differed only in a division in one terminal branch, resulting in 6 lineages, that is sufficiently short to prevent lineage sorting (Fig. 1b).

We simulated genealogies for 6 loci under  $\theta = 10.0$  and 1000 replicates. We then used STEM to compute model likelihoods for the species tree for each of these sets of simulated data under the true model as well as the false model.

## RESULTS

### *Molecular Data*

Data from 6 anonymous nuclear loci were collected and deposited in GenBank (GU197875–GU198100) and TreeBASE (SN4857). These loci averaged 625 bp and contained an average of 71 variable sites or roughly half the number identified in cytochrome *b*. Models of sequence evolution were selected for each locus using DT-ModSel (Table 2). Gene trees are characterized by the general lack of species (and subspecies) monophyly. The molecular clock was rejected using an LRT at 3 of the 6 anonymous nuclear loci (Table 2). The rejections of the molecular clock at these loci were apparently not a function of a single individual (Supplementary Table 1). Results from HyPhy indicate that there is some evidence for recombination at the loci for which the molecular clock hypothesis was rejected (Supplementary Table 2), but these findings are not supported by the

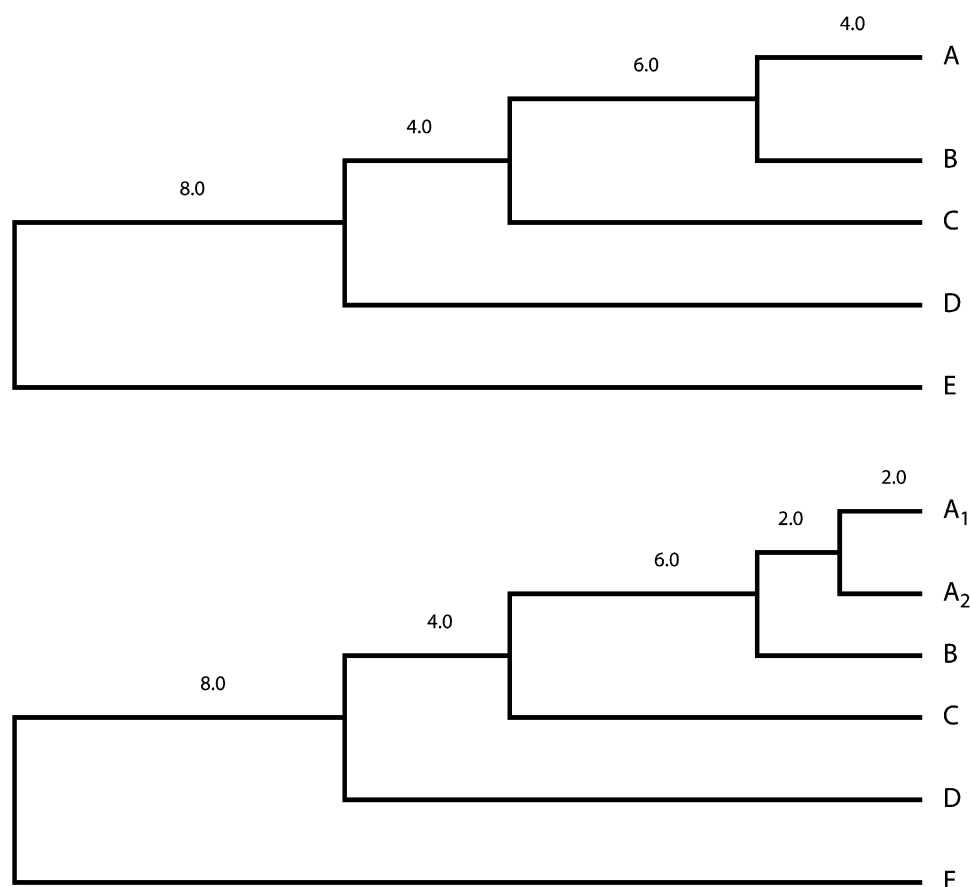


FIGURE 1. Species phylogenies used in the false lumping/splitting simulations. Branch lengths are in units of  $N$  generations, and genealogies were simulated using  $\theta = 10.0$ .

TABLE 2. Information about each locus

Locus	bp	<i>n</i>	<i>s</i>	Indels	Model	−lnL clock	−lnL unc	LRT	df
681a	732	26	53	3	K80+I	1595.86433	1579.12674	33.47518	24
681b	546	37	69	None	K81uf+I	1493.57263	1474.77693	37.5914	35
685a	605	38	80	5	K81+I+Γ	1619.8739	1595.43153	48.88474	36
734z	718	37	110	8	HKY+I	2183.66386	2135.17373	96.98026*	35
735b	627	42	73	6	TIMEf+I+Γ	1809.92806	1755.56346	108.7292*	40
735f	523	46	43	3	K80+I+Γ	1399.23408	1346.43538	105.5974*	44
cytb	741	34	134	None	HKY+Γ	2083.05541	2067.00842	32.09398	32

Note: Shown for each locus are the length (bp), the number of alleles (*n*), segregating sites (*s*), indels, the model of sequence evolution, the log-likelihoods with the molecular clock enforced and not enforced, the test statistic for the LRT, and the degrees of freedom (df).

TOPALi analyses. The HKA test identified no evidence of selection (Supplementary Table 2).

### Species Tree Estimation

The species tree estimated by MDC using Mesquite recovered a monophyletic *M. lucifugus* when subspecies were used as OTUs. However, *M. thysanodes* and *M. evotis* were paraphyletic in this species tree (Supplementary Fig. 1), and the latter was still paraphyletic when *M. thysanodes* was collapsed into a single lineage (not shown). When species were used as OTUs, *M. lucifugus* and *M. evotis* were sister taxa, with *M. keenii* sister to this clade and *M. thysanodes* basal to the other members of this clade. Nodal support values were generally low for either species tree (Supplementary Fig. 1).

Estimates of  $\theta$  were broadly similar across species and subspecies (Table 3), suggesting that this parameter does not vary appreciably across lineages. Additionally, the results of species delimitation tests (below) do not change (e.g., the rank order of models is identical and the model likelihoods are similar) depending on whether we use a value of  $\theta$  averaged across estimates from subspecies ( $\theta = 10.01$ ) or species ( $\theta = 10.12$ ). We report results using the value of  $\theta$  estimated from subspecies because the model that treats subspecies as OTUs has a better likelihood given the data than the model that treats species as OTUs, regardless of the  $\theta$  used.

TABLE 3. Per-lineage estimates of  $\theta$  from Migrate-n

Group	Low	$\theta$	High
Ave (species)	8.383	10.127	12.362
<i>M. lucifugus</i>	11.859	13.471	15.325
<i>M. evotis</i>	6.970	8.721	11.054
<i>M. thysanodes</i>	6.318	8.187	10.706
Ave (subspecies)	7.610	10.011	13.236
<i>M. lucifungus lucifugus</i>	9.652	12.284	15.468
<i>M. l. alacensis</i>	6.481	8.800	11.977
<i>M. l. carissima</i>	6.522	8.919	12.407
<i>M. l. relictus</i>	8.765	11.591	15.448
<i>M. e. chrysonotus</i>	6.630	8.460	10.877

Notes: Shown are the ML estimates of  $\theta$  (average per locus) for each species and subspecies, as well as the averages at each taxonomic level. Also shown are the upper and lower 95% confidence intervals. All estimates for any partition fall well within the confidence intervals for all other partitions. Note that we have not included estimates for partitions that sampled three or fewer individuals.

Species tree estimates (using STEM) were made using loci that were consistent with the molecular clock, as well as with all 6 loci. The species phylogeny (Fig. 2a) suggests that *M. keenii* is the basal member of the *M. lucifugus* and the western long-eared *Myotis* clade, that *M. thysanodes* is the next most basal lineage, and that *M. lucifugus* and *M. evotis* are sister taxa. As in the Mesquite MDC analysis, using described subspecies as OTUs changes the phylogeny estimate (Fig. 2b). Although the overall structure of the topology is similar to that where species are used as OTUs, both *M. evotis* and *M. lucifugus* are paraphyletic. Additionally, the overall depth of the topology is less when subspecies are used as OTUs. The topology is constant when the species tree is estimated using all loci (e.g., including the 3 that violate the molecular clock), but the timing of speciation is compressed toward the present (Fig. 2c). When only the non-clocklike loci are used, the topology matches that found in the Mesquite MDC analysis (Fig. 2d).

We conducted several BEST analyses, including 2 sets that mimicked the STEM analyses (e.g., treating species and subspecies as OTUs). BEST runs lasted a minimum of 101,000,000 generations, and samples from the posterior distribution were recorded every 1000 generations. Several methods were used to assess stationarity of Markov chains. For 3 of these 4 BEST analyses, plots of the log-likelihood suggested that the Markov chains reached stationarity within 1,000,000 generations, but the standard deviation of split frequencies for coupled runs ranged between 0.12 and 0.18 during the post-burn-in portion of the Markov chain, suggesting that they are sampling different regions of gene tree parameter space. However, the parameter of greatest interest, the topology of the species tree, did not change during this time frame, and the likelihood plots during this portion of the chain were flat. In the other runs (e.g., subspecies as OTUs with all loci), the plot of the log-likelihood increased during the course of the run and the topology of the species tree was not stable. We concluded that the posterior distribution of this run was not sampling from a stationary region and consequently do not discuss it further.

The phylogenetic tree visualization program TreeAnnotator, distributed with the BEAST software package d (Drummond and Rambaut 2007), was used to explore both nodal support and branch length in the posterior distributions of the BEST runs (Fig. 3). When all the data are used, the species phylogeny that is most

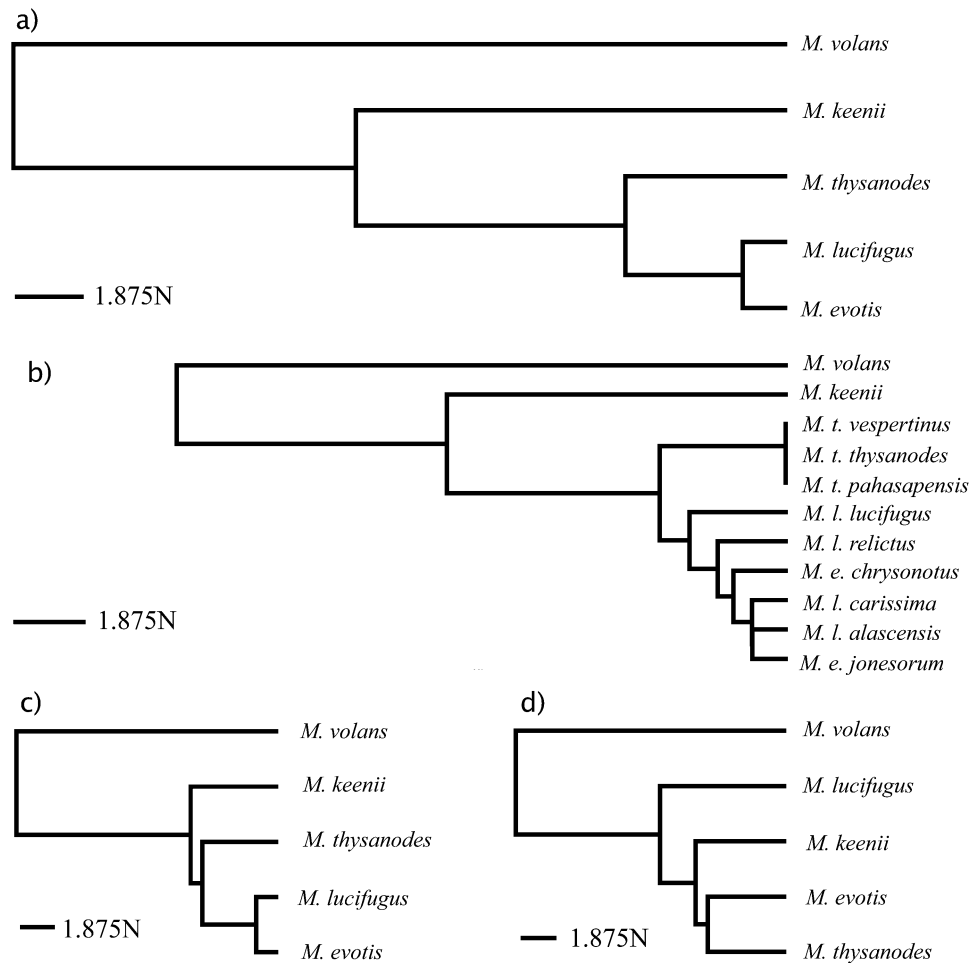


FIGURE 2. Species trees estimated using STEM. a) Shows the species tree estimated using the species as OTUs and only the clocklike loci, whereas (b) shows the species tree estimated using the subspecies as OTUs. c) Shows the species tree estimated using all loci, including those that violate the molecular clock assumption. d) Shows the topology estimated when only the non-clocklike loci are used. Note that the trees shown in (a) and (b) are drawn to equivalent scales, as are the trees shown in (c) and (d).

frequently sampled in the posterior distribution is similar to that from the Mesquite MDC analysis. *Myotis lucifugus* is basal to the clade composed of *M. keenii*, *M. evotis*, and *M. thysanodes* (Fig. 3a); nodal support is high for the node uniting *M. lucifugus* and the western long-eared *Myotis* and lower within the clade composed of the latter group. When only the clocklike loci are used, the topology changes, as does the timing of speciation, which occurs deeper in the phylogeny. *Myotis lucifugus* and *M. evotis* are sister taxa, with *M. thysanodes* basal to this clade and *M. keenii* basal to the other species (Fig. 3b). Results from the analysis that treated subspecies as OTUs suggest that the subspecies within *M. lucifugus* form a monophyletic clade, as do the subspecies that composed the western long-eared *Myotis* species (Fig. 3c). However, the latter species are not monophyletic in this analysis.

Concatenation across loci using a partitioned model in MrBayes was also used as an approach to estimating the phylogeny. Strictly interpreted, the results (Fig. 4) suggest that *M. thysanodes*, *M. evotis*, and *M. lucifugus* are

each paraphyletic, with paraphyly of the latter caused by a single sample (CO4). In general, posterior probabilities of nodes deep in the phylogeny are high, but based on the concatenated phylogeny, it is impossible to differentiate 3 plausible scenarios: 1) diversification of this clade was extremely rapid and occurred in the recent past, all shared polymorphism results from incompletely sorted ancestral polymorphism, or 2) the actual lineage boundaries do not correspond to described taxonomic groups, complicating interpretation of the concatenated phylogeny, or 3) the taxonomy is correct, but gene flow and hybridization are rampant.

#### Tests for Lineage Delimitation

The likelihoods of the species trees for 150 permutations of lineage composition were calculated by STEM, first using only clocklike loci and second using all loci with branch lengths constrained to fit the molecular clock. Because results were similar (see Supplementary Tables 3 and 4), we limit our discussion to the

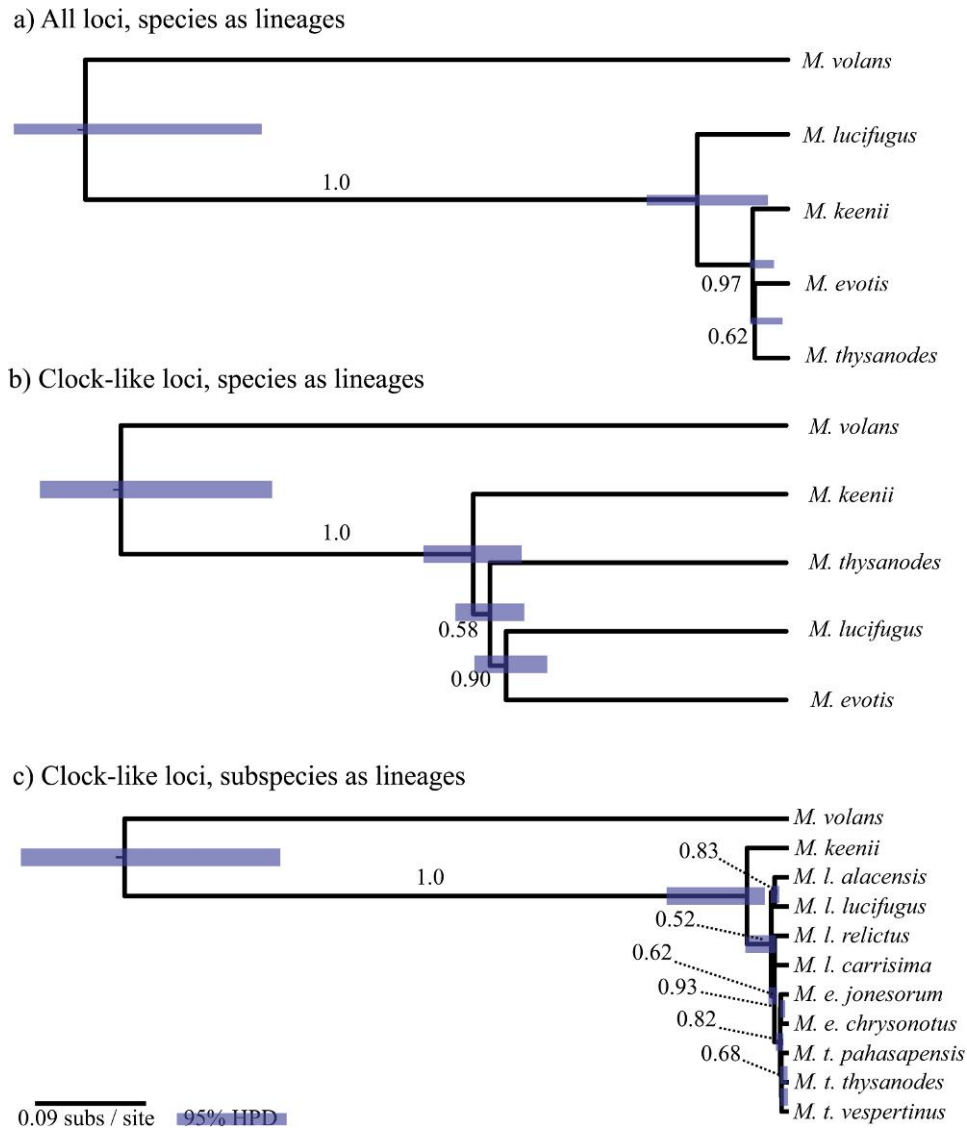


FIGURE 3. Representation of the posterior distribution of species tree space from a BEST analysis. a) Shows the species tree estimated using species as OTUs and all loci, whereas (b) shows the species tree estimated using the 3 loci that were consistent with the molecular clock. c) Shows phylogeny estimate from the clocklike loci and subspecies as OTUs. For each, the branches are the mean values of all tree in the posterior distribution, the shaded bars represent the 95% highest posterior density interval of branch lengths, and the numbers above each branch depict the Bayesian posterior probabilities of each node.

first set of calculations. We first compared a model where each described subspecies within a given species was independent with models that collapse some combination of these subspecies into lineages. For these tests, which were conducted using samples across all subspecies, we chose models with the highest overall likelihood (e.g., no constraints were placed on lineage composition in other species) and then used LRTs to assess statistical significance after correcting for multiple comparisons with a Bonferroni correction. For both *M. lucifugus* and *M. evotis*, collapsing subspecies into lineages results in a significant decrease in the likelihood of the model given the data (Table 4), suggesting that many of the described subspecies within these species are independent evolutionary lineages. Conversely,

results suggest that *M. thysanodes* is composed of a single evolutionary lineage. Bayes factors were computed for some comparisons by comparing the output of several BEST runs (Table 5). When the model that treated described subspecies as lineages was compared with the model that treated species as lineages, the Bayes factors favored the former. Similarly, this approach favored models where 2 of the 3 subspecies within *M. thysanodes* were collapsed into a single lineage but not the model where all 3 subspecies were collapsed. In all cases, relative support for the preferred model was not strong.

We also evaluated all models using AIC and metrics borrowed from information theory (summarized in Table 6). Four models have high likelihood given the



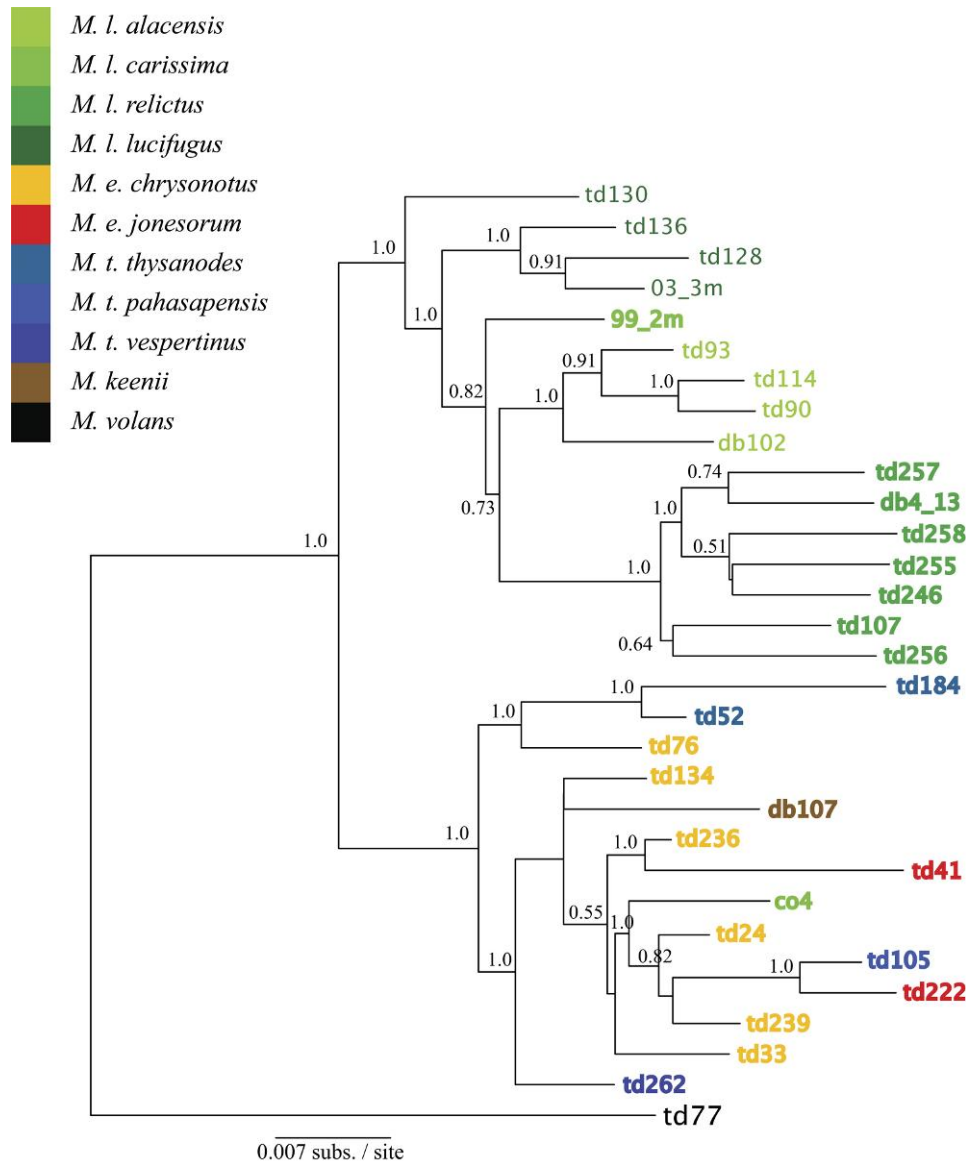


FIGURE 4. Phylogeny estimated from the concatenated data. Unlike the other methods used to estimate phylogeny, individuals serve as the OTUs in this phylogeny. We randomly chose one of the phased alleles from the nuclear loci and combined them with the cytochrome *b* data, we then repeated the analysis with the other set of alleles. Shown is the tree from one run, the other was broadly similar. Samples are color-coded to subspecies using the colors shown to the left of the phylogeny.

data and are quite similar to one another. In each of these models, *M. thysanodes* forms a single or 2 evolutionary lineages and the subspecies within *M. lucifugus* and *M. evotis* are independent evolutionary lineages. Together, these models represent >96% of the total model probabilities, indicating that there is little support for the other 146 models. These results are similar when all loci are used in the analysis (Supplementary Table 4). The results from the LRTs and the information-theoretic analysis are consistent in that they suggest that most of the currently described species within the *M. lucifugus*/western long-eared *Myotis* clade contain more than one evolutionarily independent lineage.

#### *Effects of Incorrect Assignment of Individuals to Lineages on Species Tree Estimation*

Using STEM, we explored the effects of false lumping and false splitting by estimating the species tree using genealogies simulated under the models shown in Figure 1. STEM is very accurate when the assignment of individuals to lineages matches the simulation conditions as the species tree averaged across 1000 replicates matches the simulation topology closely (Fig. 5a,b). When alleles from separate evolutionary lineages are falsely lumped, divergence time is overestimated, although the topology is correct. False splitting does not result in a similar bias as false lineages essentially coalesce at the tip of the phylogeny. These effects are

TABLE 4. Delimitation using LRTs

Species	Lineage composition	−lnL	2(Δ−lnL)
<i>M. thysanodes</i>	Each spp. separate	570.533	
	<i>Mtp_Mtv_Mtt</i>	570.533	0.000
	<i>Mtt_Mtp</i>	570.533	0.000
	<i>Mtt_Mtv</i>	570.533	0.000
	<i>Mtp_Mtv</i>	573.375	5.686
<i>M. evotis</i>	<i>Mec, Mej</i>	570.533	
	<i>Mec_Mej</i>	579.581	18.097*
<i>M. lucifugus</i>	Each spp. separate	570.533	
	<i>Mla_Mlc_Mll_Mlr</i>	648.174	155.282*
	<i>Mla_Mlc_Mll</i>	620.935	100.804*
	<i>Mla_Mlc_Mlr</i>	588.732	36.398*
	<i>Mla_Mll_Mlr</i>	637.752	134.440*
	<i>Mlc_Mll_Mlr</i>	635.230	129.395*
	<i>Mla_Mlc</i>	573.375	5.686*
	<i>Mla_Mll</i>	613.120	85.175*
	<i>Mla_Mlr</i>	583.284	25.504*
	<i>Mlc_Mll</i>	611.458	81.850*
	<i>Mlc_Mlr</i>	582.416	23.767*
	<i>Mll_Mlr</i>	602.181	63.298*
	<i>Mla_Mlc, Mll_Mlr</i>	605.059	69.052*
	<i>Mla_Mll, Mlc_Mlr</i>	630.831	120.597*
	<i>Mla_Mlr, Mlc_Mll</i>	630.017	118.969*

Notes: Shown for each species are the comparisons, the highest likelihood given a particular model, and the test statistic. Within each species, we used a Bonferroni correction for multiple comparisons. Significant values are shown with an “\*”. The first letter of the genus, species, and subspecies names are used to abbreviate each subspecies (e.g., *Myotis evotis jonesorum*).

manifested in the likelihood scores; results suggest that falsely lumping members of 2 evolutionary lineages will decrease the model likelihoods, whereas falsely splitting the members of a single lineage will not. Clearly, users of STEM should favor splitting to lumping when the boundaries of the lineages are in question because false lumping appears to have important consequences to phylogeny estimation, whereas false splitting does not.

## DISCUSSION

Systematic investigations that use genetic data to evaluate existing taxonomic descriptions are often complicated by findings of poly- or paraphyly in described species (Funk and Omland 2003). Although many recent investigations have documented cases of appreciable intraspecific genetic divergence (Hendrixson and Bond 2004; Olson et al. 2004; Camargo et al. 2006; Kozak et al. 2006; Starrett and Hedin 2007), it is not yet clear how best to use molecular data to test existing taxonomic

designations. One approach, which we have adopted here, proceeds by evaluating models of lineage composition under a phylogenetic framework that implements a coalescent model. Because population-level processes such as the loss of ancestral polymorphism due to genetic drift are ubiquitous at the initial stages of diversification, we suspect that accurate estimates of phylogeny at the shallowest levels of divergence will require phylogenetic methods that incorporate coalescent models. Although acceptance of these models is growing among systematists (Edwards 2009), it is not commonly appreciated that the application of methods such as MDC, STEM, and BEST requires an active consideration of the boundaries of lineages because the underlying coalescent models assume that each OTU consists of a single metapopulation lineage. Here, we make the first attempt to test models of lineage composition in a complex empirical system using a probabilistic framework.

Two approaches to lineage delimitation were used. In the first, LRTs were used to test models that lumped described subspecies into aggregated lineages. In *M. evotis*, where individuals from 2 subspecies were sampled, this approach required a single comparison, and the LRT suggests that we can reject the model where these subspecies constitute a single lineage. However, in species where there are more than 2 subspecies (here *M. thysanodes* and *M. lucifugus*), models representing all combinations of lineage composition must be compared. Although our results were internally consistent in that we do not identify a subspecies as independent in one comparison and nonindependent in another, such results are possible. Therefore, we developed a second approach to delimitation that proceeds by exhaustively calculating the probability of hierarchical combinations of lineage composition, computing AIC scores for each, and calculating the relative likelihood of the models given the data. Information-theoretic results provide both an absolute ranking of the models and an evaluation of the relative statistical support of the best models and are thus easier to interpret in complex systems such as *M. lucifugus* and the western long-eared *Myotis*. For example, in our *Myotis* data, we find that models treating the majority of described subspecies as evolutionary lineages have the highest likelihood, regardless of methodological details such as the assumed value of  $\theta$  or the inclusion of loci that violate the molecular clock.

We assessed this strength of statistical support of a given model of lineage composition by calculating several statistics. Model probabilities ( $w_j$ ), which are

TABLE 5. Results of species delimitation using BEST and the Bayes factor

Lineage Composition	−lnL	BF	Interpretation
<i>Mla, Mlc, Mll, Mlr, Mec, Mej, Mtp, Mtt, Mtv</i>	12502.211		Por
<i>Mla_Mlc_Mll_Mlr, Mec_Mej, Mtp_Mtt_Mtv</i>	12857.375	0.9724	Modest support subspecies as lineages
<i>Mla, Mlc, Mll, Mlr, Mec, Mej, Mtp_Mtv, Mtt</i>	11830.384	1.0568	Modest support for <i>Mtp_Mtv</i> forming a single lineage
<i>Mla, Mlc, Mll, Mlr, Mec, Mej, Mtt_Mtp, Mtv</i>	11851.727	1.0549	Modest support for <i>Mtt_Mtp</i> forming a single lineage
<i>Mla, Mlc, Mll, Mlr, Mec, Mej, Mtt_Mtv_Mtp</i>	12598.799	0.9923	Modest support for <i>M. thysanodes</i> (inclusive)

Note: Shown are the details of the lineages combined for a given BEST analysis, the mean log-likelihood, the Bayes factor (BF), and the interpretation of these values.

TABLE 6. Summary of information-theoretic assessment of species limits

Lineage composition	−lnL	<i>k</i>	AIC	Δ <i>i</i>	<i>L</i> (Model data)	<i>w<sub>i</sub></i>
<b>Clocklike loci</b>						
<i>Mtp_Mtt_Mtv, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−570.533	8	1157.065	0.000	1.000	0.685
<i>Mtp_Mtv, Mtt, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−570.533	9	1159.065	2.000	0.135	0.093
<i>Mtt_Mtp, Mtv, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−570.533	9	1159.065	2.000	0.135	0.093
<i>Mtt_Mtv, Mtp, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−570.533	9	1159.065	2.000	0.135	0.093
<i>Mtp_Mtt_Mtv, Mla_Mlc, Mll, Mlr, Mej, Mec</i>	−573.375	7	1160.751	3.686	0.025	0.017
<i>Mtp, Mtt, Mtv, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−570.533	10	1161.065	4.000	0.018	0.013
<i>Mtp_Mtv, Mtt, Mla_Mlc, Mll, Mlr, Mej, Mec</i>	−573.375	8	1162.751	5.686	0.003	0.002
<i>Mtt_Mtp, Mtv, Mla_Mlc, Mll, Mlr, Mej, Mec</i>	−573.375	8	1162.751	5.686	0.003	0.002
<i>Mtt_Mtv, Mtp, Mla_Mlc, Mll, Mlr, Mej, Mec</i>	−573.375	8	1162.751	5.686	0.003	0.002
<i>Mtp, Mtt, Mtv, Mla_Mlc, Mll, Mlr, Mej, Mec</i>	−573.375	9	1164.751	7.686	0.000	0.000
<b>All loci</b>						
<i>Mtp_Mtt_Mtv, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−1242.720	8	2501.440	0.000	1.000	0.775
<i>Mtt_Mtv, Mtp, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−1242.716	9	2503.432	1.992	0.136	0.106
<i>Mtt_Mtp, Mtv, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−1242.720	9	2503.440	2.000	0.135	0.105
<i>Mtp, Mtt, Mtv, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−1242.720	10	2505.440	4.000	0.018	0.014
<i>Mtp_Mtt_Mtv, Mla_Mll, Mlc, Mlr, Mej, Mec</i>	−1248.517	7	2511.035	9.595	0.000	0.000
<i>Mtt_Mtv, Mtp, Mla_Mll, Mlc, Mlr, Mej, Mec</i>	−1248.513	8	2513.027	11.586	0.000	0.000
<i>Mtt_Mtp, Mtv, Mla_Mll, Mlc, Mlr, Mej, Mec</i>	−1248.517	8	2513.035	11.594	0.000	0.000
<i>Mtp, Mtt, Mtv, Mla_Mll, Mlc, Mlr, Mej, Mec</i>	−1248.517	9	2515.035	13.594	0.000	0.000
<i>Mtp_Mtv, Mtt, Mla, Mlc, Mll, Mlr, Mej, Mec</i>	−1249.002	9	2516.005	14.564	0.000	0.000
<i>Mtp_Mtt_Mtv, Mlc_Mlr, Mla, Mll, Mej, Mec</i>	−1251.395	7	2516.790	15.349	0.000	0.000

Note: Columns (from left) show the set used as OTUs (using the abbreviations shown in Table 1), the log-likelihood of the model given the data, the number of parameters, the AIC, AIC differences (Δ*i*), relative likelihood of the model given the data, and the model probabilities.

essentially relative likelihoods of the model given the data (Anderson 2008), are calculated by first computing the difference in AIC between the best model and model *i*, then summing the relative likelihoods of all models given the data, and finally determining what proportion of the likelihood of all models given the data is represented by model *i*. Because the AIC difference (Δ*i*) is a measure of the difference in Kullback–Leibler divergence between the models, these metrics essentially represent how much information is lost when model *i* is used to represent reality rather than the best model. In the case of our *Myotis* data, a substantial amount of information loss occurs when evolutionary lineages (here subspecies) are lumped without biological justification. For example, the Δ*i* between the best model and the model where described species are treated as independently evolving lineages is >180 AIC units, suggesting that the model consistent with the existing taxonomy is extremely poor. Model probabilities provide another way to assess the relative merits of a set of models; here, the 6 best models (which all treat subspecies within *M. lucifugus* and *M. evotis* as independent) account for over 99% of the total model probabilities. Clearly, there is little support in our data for considering *M. lucifugus* and *M. evotis* as 2 species that each consists of a single evolutionary lineage.

Our empirical results also suggest that gene trees from loci that violate the molecular clock may bias estimates of species phylogeny using STEM. Qualitatively, adding these loci compresses the temporal divergence toward the present (Fig. 2c), a behavior that is expected for a method (STEM) that is based on divergence time estimation. It is interesting to note that BEST includes method for converting nonultrametric gene trees (which are proposed during the Markov chain Monte Carlo [MCMC] search of gene tree space) into ultrametric

gene trees used to calculate the posterior density of gene trees given the species tree. To date, the effects of this approximation have not been evaluated (Rannala and Yang 2008), and in general these sorts of evaluations are computationally difficult in programs such as BEST that implement MCMC methods. However, when we conducted BEST runs using all loci, we observed a compression of the temporal divergence toward the present that was qualitatively similar to the behavior exhibited by STEM (cf. Figs. 2c–3b). Obviously, these comparisons are difficult to make because clocklike loci are a subset of the total. We also observed that the species phylogeny estimated from all loci differs from one estimated from a subset of the loci in the BEST analysis. STEM and BEST complement one another in 2 important ways. STEM assumes that gene trees are estimated without error, an assumption that is almost certainly violated for empirical systems. BEST uses separate Markov chains to estimate the gene trees and the species phylogeny. Consequently, the posterior density of species tree space is integrated over the uncertainty in gene tree space. However, Markov chains are computationally intensive to a degree that makes detailed exploration via simulation difficult (but see Liu et al. 2008). STEM is computationally inexpensive by comparison, which allows for exploration of factors related to lineage delimitation such as the number and composition of lineages. These contrasts suggest to us that neither of these programs is likely to be as effective independently as they are when they are used concurrently.

Although our results suggest that a coupled approach to delimitation and phylogeny estimation can be used to elucidate evolutionary history in complex radiations such as in the *M. lucifugus*/western long-eared *Myotis* clade, we have several concerns about our approach. One of the most significant is the assumption that our

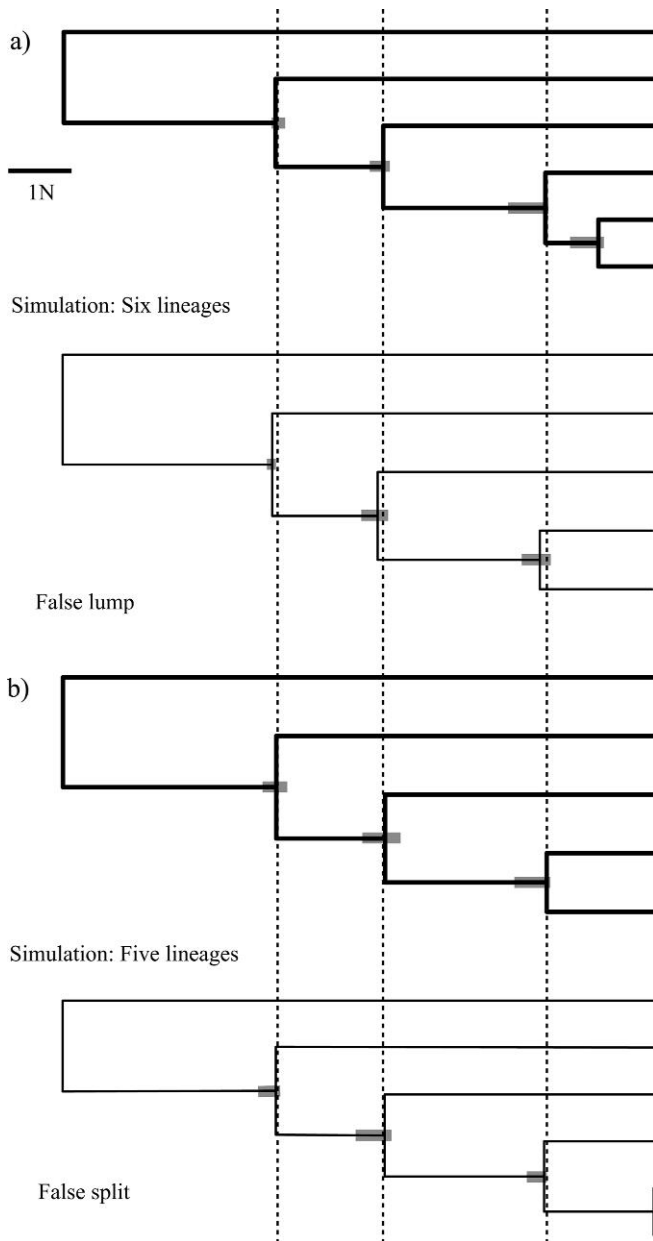


FIGURE 5. Example results from simulation exploration of species delimitation using STEM. a) Shows the average of species tree estimates made using 10 loci with 6 lineages as OTUs (phylogeny with thick lines) and the effects of falsely lumping 2 lineages together. b) Shows the effects of falsely splitting a single lineage. In all cases, the topology is consistent, and the species tree estimated using the correct assignment of individuals to lineages is essentially identical to that used for the simulations. Shaded boxes represent the 95% confidence interval of branch length for the species tree estimates.

gene trees are estimated without error because phylogenetic error in gene tree estimation is associated with a lack of resolution in the species tree estimates (Huang and Knowles 2009). Nodal support values for each estimated gene tree (Supplementary Fig. 1) are relatively high, which indicate to us that our gene tree estimates are relatively accurate and moreover provide us with

confidence that the evolutionary lineages that we have delimited within the *M. lucifugus*/western long-eared *Myotis* clade reflect actual evolutionary units. But we cannot be certain that our estimates of the genealogies are accurate, this is a shortcoming inherent to all approaches that estimate species trees from gene trees. The other significant concern is the assumption that the shared polymorphism is entirely the result of the incomplete sorting of ancestral polymorphism. We explored this concern by estimating gene flow, under the assumption that it was the other likely source for shared polymorphism in this system. Two dramatically different coalescent methods were used, and this exploration highlights both the difficulties and the possibilities of working at the interface between phylogenetics and population genetics.

#### Assessing the Prevalence of Gene Flow

All methods for estimating species trees, and consequently the approaches to delimitation outlined above, assume that the shared polymorphism results from incompletely sorted ancestral polymorphism. Gene flow can decrease the accuracy of species tree estimates under some conditions (Eckert and Carstens 2008) and it is reasonable to expect that it could do the same for delimitation. However, just as estimates of species phylogeny may be complicated by gene flow, estimating gene flow in empirical systems can be complicated by the divergence among populations (Wakeley 2000). Generally, there are 2 approaches that one can take when designing a coalescent-based method to estimate gene flow. If one assumes an equilibrium model, gene flow can be estimated without considering phylogeny; we used one such model (Migrate-n) to estimate migration rates among the 6 described subspecies within *M. lucifugus* and *M. evotis*. Our results indicate that migration rates can be appreciable and differ by an order of magnitude across pairwise comparisons (Supplementary Table 5). However, doing so ignores the phylogenetic divergence within this clade, and thus we consider it to be an inappropriate application of the method. When nonequilibrium methods are used to estimate gene flow, the estimates of gene flow are much lower. For example, IMA (Hey and Nielsen 2007) was used to estimate gene flow,  $\theta = 4N_e\mu$ , and population divergence for several pairwise comparisons of subspecies. An advantage of IMA over its precursors is the capacity to evaluate the strength of support for a set of demographic models that reduce the parameters of the full model (e.g.,  $\theta_A \theta_1 \theta_2 m_{12} m_{21} \tau$ ) in various ways using information theory (Carstens et al. 2009). Briefly, this method quantifies the probability of multiple models given the data thereby allowing researchers to identify the subset of parameters in the full model that represent important biological processes. For example, if all shared polymorphism is best explained by incomplete lineage sorting rather than migration, the selected models would not include migration parameters (e.g., it would only include  $\theta_A \theta_1 \theta_2 t$ ). Results (Supplementary Table 6) suggest



that gene flow is not an important parameter to include in the demographic model for 3 comparisons (*M. l. alascensis*–*M. l. carissima*; *M. e. jonesorum*–*M. l. carissima*; *M. l. alascensis*–*M. e. jonesorum*) but is necessary for one (*M. e. chrysonotus*–*M. l. alascensis*). Note that we do not have the level of phylogeographic sampling required to thoroughly investigate this question, but the approach outlined above may represent a useful method for exploring this question.

#### *Evolution of M. lucifugus and the Western Long-eared Myotis*

Although the phylogeny estimates reported here (Figs. 2–4) are preliminary and will likely change as data are added, they contain valuable information regarding the timing of divergence. Branch lengths in Figure 2 are not shown in the customary substitutions/site but rather in coalescent units equal to  $N$  generations. If we assume the rate of mutation used by Stadelmann et al. (2007) in their molecular dating of the *Myotis* radiation, our estimate of  $N_e$  is  $\sim 300,000$  (averaged across lineages). Generation lengths are in the range of 3–5 years, so one coalescent unit is equivalent to 1–1.5 My of divergence. This implies that most of the diversification within the *M. lucifugus*/western long-eared *Myotis* clade occurred during the Pleistocene.

The morphological flexibility of this recent radiation is striking. *Myotis lucifugus* (inclusive) is a foraging generalist, whereas western long-eared *Myotis* species specialize on moth prey and possess morphological adaptations associated with that specialization. Western long-eared *Myotis* species are found primarily in coniferous forested regions of western North America, suggesting that differentiation in this group could have been influenced by expansion and fragmentation of suitable habitat in response to glacial cycling. Future sampling will focus on thorough sampling of described subspecies throughout their ranges to further refine the delimitation of lineages in this complex group and better understand their biogeography and rapid morphological divergence.

#### *Future Directions for Species Delimitation*

Methods for species phylogeny estimation will continue to improve, particularly with the expansion of phylogenetic methods to relax the assumption that all discord between gene trees and species trees is caused by genetic drift (Kubatko 2009) and the expansion of the isolation-with-migration model to include more than 2 lineages (Hey 2010). We also envision several ways that delimitation methods could be improved. Assuming that previously described sets of samples were used, delimitation efforts could proceed by more thoroughly exploring possible combinations. For example, rather than exhaustively calculating the likelihood of hierarchical models of subspecies, as we have done here, one could proceed by calculating all combinations of all subspecies either in an exhaustive manner or heuristically.

We restricted our analyses to hierarchical combinations in part because there are over 100,000 combinations of 10 subspecies, a number too large to exhaustively evaluate. However, it is worth noting that several combinations where *M. e. jonesorum* is grouped with either or both of *M. l. alascensis* or *M. l. carissima*, as suggested in Figure 2a, have likelihoods that are relatively poor and would account for less than 1% of the total model probabilities. Alternatively, one could parameterize lineage membership and use MCMC methods to approximate the posterior probability of lineage composition in addition to the species phylogeny. Finally, an exploration of the similarities and differences between phylogenetic approaches to delimitation and methods such as Structure (Pritchard et al. 2000) would be of great interest. Regardless of which methods are employed, we are optimistic that these advances will allow systematists to delimit evolutionary lineages at the initial stages of divergence and identify the ecological and environmental forces that promoted diversification.

#### SUPPLEMENTARY MATERIAL

Supplementary material can be found at: <http://www.sysbio.oxfordjournals.org/>.

#### FUNDING

University of Michigan Ecology and Evolutionary Biology Department and Museum of Zoology, American Wildlife Foundation, American Society of Mammals, and Sigma Xi.

#### ACKNOWLEDGMENTS

For their help in fieldwork and obtaining samples we thank Dr Kevin Castle (National Park Service), Doug Burles (Gwaii Hanaas National Park and Haida Heritage Site), Kristi Jacques (Montana Fish, Wildlife, and Parks), and Kirk Navo (Colorado Division of Wildlife). We thank J. Carstens for assistance collecting data, and D. Triant and members of the Carstens lab for comments on the many drafts of this article. We thank M. Hedin and several anonymous reviewers for comments and suggestions that greatly improved this article.

#### REFERENCES

- Akaike H. 1973. Information theory as an extension of the maximum likelihood principle. In: Petrov B.N., Csaki F., editors. Second International Symposium on Information Theory. Budapest (Hungary): Akademiai Kiado. p. 267–281.
- Anderson D.R. 2008. Model based inference in the life sciences. New York: Springer.
- Ané C., Larget B., Baum D.A., Smith S.D., Rokas A. 2007. Bayesian estimation of concordance among gene trees. *Mol. Biol. Evol.* 24: 412–426.
- Baker R.H., Stains H.J. 1955. A new long-eared *Myotis* (*Myotis evotis*) from northeastern Mexico. *Univ. Kansas Mus. Nat. Hist.* 9:81–84.
- Berli P., Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in  $n$  subpopulations

- by using a coalescent approach. *Proc. Natl. Acad. Sci. USA.* 98: 4563–4568.
- Belfiore N.M., Liu L., Moritz C. 2008. Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Geomyidae). *Syst. Biol.* 57:294–310.
- Brumfield R.T., Beerli P., Nickerson D.A., Edwards S.V. 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol. Evol.* 18:249–256.
- Brumfield R.T., Liu L., Lum D.E., Edwards S.V. 2008. Comparison of species tree methods for reconstructing the phylogeny of bearded manakins (Aves: Pipridae, *Manacus*) from multilocus sequences data. *Syst. Biol.* 57:719–731.
- Burnham K.P., Anderson D.R. 1998. Model selection and inference: a practical information-theoretic approach. New York: Springer.
- Camargo A., De Sa R.O., Heyer W.R. 2006. Phylogenetic analyses of mtDNA sequences reveal three cryptic lineages in the widespread neotropical frog *Leptodactylus fuscus* (Schneider, 1799) (Anura, Leptodactylidae). *Biol. J. Linn. Soc.* 87:325–341.
- Carstens B.C., Knowles L.L. 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Syst. Biol.* 56:400–411.
- Carstens B.C., Reid N.M., Stoute H.N. 2009. An information theoretic approach to phylogeography. *Mol. Ecol.* 18:4270–4282.
- Creer S. 2007. Choosing and using introns in molecular phylogenetics. *Evol. Bioinf.* 3:99–108.
- Degnan J.H., Rosenberg N.A. 2006. Discordance of species trees with their most likely gene trees. *PLoS Genet.* 2:0762.
- Degnan J.H., Salter L.S. 2005. Gene tree distributions under the coalescent process. *Evolution.* 9:24–37.
- de Querioz K. 2007. Species concepts and species delimitation. *Syst. Biol.* 56:879–886.
- Dewey T.A. 2006. Systematics and phylogeography of North American *Myotis* [dissertation]. [Ann Arbor (MI)]: University of Michigan.
- Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:e214.
- Eckert A.J., Carstens B.C. 2008. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Mol. Phylogenet. Evol.* 49:832–842.
- Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? *Evolution.* 63:1–19.
- Edwards S.V., Liu L., Pearl D.K. 2007. High-resolution species trees without concatenation. *Proc. Natl. Acad. Sci. USA.* 104:5936–5941.
- Funk D.J., Omland K.E. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Syst.* 34:397–423.
- Harrison R.G. 1998. Linking evolutionary pattern and process: the relevance of species concepts for the study of speciation. New York: Oxford University Press.
- Hendrixson B.E., Bond J.E. 2004. Testing species boundaries in the *Antrodiaetus unicolor* complex (Araneae: Mygalomorphae: Antrodiaetidae): “paraphyly” and cryptic diversity. *Mol. Phylogenet. Evol.* 36:405–416.
- Hey J. 2010. Isolation with migration models for more than two populations. *Mol. Biol. Evol.* 27:905–920.
- Hey J., Nielsen R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. USA.* 104:2785–2790.
- Huang H., Knowles L.L. 2009. What is the danger of the anomaly zone for empirical phylogenetics? *Syst. Biol.* 58:527–536.
- Hudson R.R. 1991. Gene genealogies and the coalescent process. In: Futuyma D., Antonovics J., editors. *Oxford surveys in evolutionary biology*. New York: Oxford University Press. p. 1–44.
- Hudson R.R. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics.* 18:337–338.
- Hudson R.R., Coyne J.A. 2002. Mathematical consequences of the genealogical species concept. *Evolution.* 56:1557–1565.
- Hudson R.R., Kreitman M., Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics.* 116:153–159.
- Huelsenbeck J.P., Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* 17:754–755.
- Jukes T.H., Cantor C.R. 1969. Evolution of protein molecules. In: Munro M.N., editor. *Mammalian protein metabolism*. New York: Academic Press. pp. 21–132.
- Kass R.E., Raftery A.E. 1995. Bayes factors. *J. Am. Stat. Assoc.* 90:773–795.
- Knowles L.L., Carstens B.C. 2007. Delimiting species without monophyletic gene trees. *Syst. Biol.* 56:887–895.
- Knowles L.L., Carstens B.C., Keat M.L. 2007. Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Curr. Biol.* 17:940–946.
- Kolaczowski B., Thornton J.W. 2006. Is there a star tree paradox? *Mol. Biol. Evol.* 23:1819–1823.
- Koopman K.F. 1963. The identity of bats (genus *Myotis*) collected in Arizona by Miller, Price, and Condit in 1894. *Am. Mus. Novit.* 2140:1–3.
- Kosakovsky-Pond S.L., Frost S.D.W., Muse S.V. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics.* 21:676–679.
- Kosakovsky-Pond S.L., Posada D., Gravenor M.B., Woelk C.H., Frost S.D.W. 2006. GARD: a genetic algorithm for recombination detection. *Bioinformatics.* 22:3096–3098.
- Kozak K.H., Blaine R.A., Larson A. 2006. Gene lineages and eastern North American paleodrainage basins: phylogeography and speciation in salamanders of the *Eurycea bislineata* species complex. *Mol. Ecol.* 15:191–207.
- Kubatko L.S. 2009. Identifying hybridization events in the presence of coalescence via model selection. *Syst. Biol.* 58:478–488.
- Kubatko L.S., Carstens B.C., Knowles L.L. 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics.* 25:971–973.
- Kubatko L.S., Degnan J.H. 2009. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* 56: 17–24.
- Kullback S. 1959. Information theory and statistics. New York: John Wiley.
- Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. 2007. ClustalW2 and ClustalX version 2. *Bioinformatics.* 23:2947–2948.
- Linnen C.R., Farrell B.D. 2008. Comparison of methods for species-tree inference in the Sawfly genus *Neodiprion* (Hymenoptera: Diprionidae). *Syst. Biol.* 57:876–890.
- Liu L., Pearl D.K. 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of species phylogeny using estimated gene tree distributions. *Syst. Biol.* 56:504–505.
- Liu L., Pearl D.K., Brumfield R.T., Edwards S.V. 2008. Estimating species trees using multiple-allele DNA sequence data. *Evolution.* 62:2080–2091.
- Maddison W.P. 1997. Gene trees in species trees. *Syst. Biol.* 1997: 523–536.
- Maddison W.P., Knowles L.L. 2006. Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* 55:21–30.
- Maddison W.P., Maddison D.R. 2004. Mesquite: a modular system for evolutionary analysis. Version 1.01. Available from: <http://mesquiteproject.org>.
- McCormack J.E., Huang H., Knowles L.L. 2009. Breaking them down and building them up: evaluating the accuracy of maximum-likelihood estimates of species trees. *Syst. Biol.* 58:501–508.
- Milne I., Wright F., Rowe G., Marshall D.F., Husmeier D., McGuire G. 2004. TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics.* 20:1806–1807.
- Minin V., Abdo Z., Joyce P., Sullivan J. 2003. Performance-based selection of likelihood models for phylogeny estimation. *Syst. Biol.* 52:674–683.
- Mossel E., Vigoda E. 2005. Phylogenetic MCMC algorithms are misleading on mixture of trees. *Science.* 309:2207–2209.
- Oliver J.C. 2008. AUGUST: inferring species trees while accounting for gene tree uncertainty. *Bioinformatics.* 24:2932.
- Olson L.E., Goodman S.M., Yoder A.D. 2004. Illumination of cryptic species boundaries in long-tailed shrew tenrecs (Mammalia: Tenrecidae; *Microgale*), with new insights into geographic variation and distributional constraints. *Biol. J. Linn. Soc.* 83:1–22.

- Page R.D.M., Sullivan J. 2008. The expanding contributions of Systematic Biology. *Syst. Biol.* 57:1–3.
- Palumbi S.R., Baker C.S. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of Humpback whales. *Mol. Biol. Evol.* 11:426–435.
- Pellegrino K.C.M., Rodrigues M.T., Waite A.N., Morando M., Yassuda Y.Y., Sites J.L. 2005. Phylogeography and species limits in the *Gymnodactylus darwini* complex (Gekkonidae, Squamata): genetic structure coincides with river systems in the Brazilian Atlantic Forest. *Biol. J. Linn. Soc.* 85:13–26.
- Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155: 945–959.
- Pritham E.J., Freschotte C. 2007. Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc. Natl. Acad. Sci. USA*. 104:1895–1900.
- Rannala B., Yang Z. 2008. Phylogenetic inference using whole genomes. *Annu. Rev. Genom. Hum. Genet.* 9:217–231.
- Ray D.A., Paga M.L., Thompson M.L., Stevens R.D. 2007. Bats with hATS: evidence for recent DNA transposon activity in the genus *Myotis*. *Mol. Biol. Evol.* 24:632–639.
- Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R. 2000. Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert biota. *Proc. Natl. Acad. Sci. USA*. 97:14438–14443.
- Ronquist F., Huelsenbeck J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 19: 1572–1573.
- Rosenberg N.A., Tao R. 2008. Discordance of species trees with their most likely gene trees: the case of five taxa. *Syst. Biol.* 57:131–140.
- Rozas J., Rozas R. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*. 15:174–175.
- Rozen S., Skaletsky H.J. 2000. PRIMER<sup>3</sup> on the WWW for general users and biologist programmers. In: Krawetz S., Misener S., editors. *Bioinformatics methods and protocols: methods in molecular biology*. Totowa (NJ): Humana Press. pp. 366–386.
- Ruedi M., Mayer F. 2001. Molecular systematics of bats of the genus *Myotis* (Vespertilionidae) suggests deterministic ecomorphological convergences. *Mol. Phylogenet. Evol.* 21:436–448.
- Shaw A.J. 2000. Molecular phylogeography and cryptic speciation in the mosses, *Mielichhoferia elongata* and *M. mielichhoferiana* (Bryaceae). *Mol. Ecol.* 9:595–608.
- Sites J.W. Jr., Marshall J.C. 2004. Operational criteria for delimiting species. *Annu. Rev. Ecol. Syst.* 35:199–227.
- Stadelmann B., Lin L.-K., Kunz T.H., Ruedi M. 2007. Molecular Phylogeny of new world *Myotis* (Chiroptera, Vespertilionidae) inferred from mitochondrial and nuclear DNA genes. *Mol. Phylogenet. Evol.* 43:32–48.
- Starrett J., Hedin M. 2007. Multilocus genealogies reveal multiple cryptic species and biogeographic complexity in the California turret spider *Antridiaetis riversi* (Mygalomorphae). *Mol. Ecol.* 16: 583–604.
- Stephens M., Smith N.J., Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *Genetics*. 68: 978–989.
- Swofford D.L. 2002. PAUP\*: phylogenetic analysis using parsimony (\*and other methods). Version 4. Sunderland (MA): Sinauer Associates.
- Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics*. 105:437–460.
- Takahata N., Nei M. 1985. Gene genealogy and variance of interpopulation nucleotide differences. *Genetics*. 110:325–344.
- Templeton A.R. 2001. Using phylogeographic analyses of gene trees to test species status and process. *Mol. Ecol.* 10:779–791.
- Wakeley J. 2000. The effects of subdivision on the genetic divergence of populations and species. *Evolution*. 54:1092–1101.
- Wiens J.J. 2007. Species delimitation: new approaches for discovering diversity. *Syst. Biol.* 56:875–878.