



Incongruence in the pattern and timing of intra-specific diversification in bronze frogs and bullfrogs (Ranidae)

James D. Austin^{a,b,*}, Kelly R. Zamudio^c

^aWildlife Ecology & Conservation, 110 Newins-Ziegler Hall, P.O. Box 110430, University of Florida, Gainesville, Florida 32611, USA

^bSchool of Forest Resources and Conservation, University of Florida, 7922 NW 71st Street, Gainesville, Florida 32653, USA

^cEcology & Evolutionary Biology, E209 Corson Hall, Cornell University, Ithaca, NY 14853, USA

ARTICLE INFO

Article history:

Received 1 January 2008

Revised 14 May 2008

Accepted 18 June 2008

Available online 24 June 2008

Keywords:

Coalescent

Comparative phylogeography

Cytochrome *b*

Evolution

ND2

Parametric bootstrap

Selection

ABSTRACT

We compare patterns of lineage divergence in mitochondrial DNA (mtDNA) sequences of two protein-encoding mitochondrial genes (cyt *b* and ND2) in two ecologically similar, co-distributed, and closely related ranid frogs (*Rana clamitans* and *Rana catesbeiana*), that are geographically widespread, and frequently syntopic. We identified three lineages in *R. clamitans*, separated by 0.5% to 2.1% net corrected sequence divergence, comparable to two *R. catesbeiana* lineages separated by 0.6%. The geographic pattern of lineage distribution differed notably between the two species. In *R. clamitans*, we found a Coastal Plain-Appalachian (CPA) lineage restricted to south and east of the Appalachian Mountains and a widespread lineage that encompassing nearly all the sampled range. A third distinct and divergent lineage was detected in one location in the southwest portion of the range (Louisiana). This pattern contrasts with the east-west pattern in *R. catesbeiana*, and reflects possible differences in refugial dynamics and patterns of range expansion. Although both species have undergone range expansion and population growth, coalescent reconstruction of N_e reflects larger lineages but more recent divergence in *R. clamitans* relative to *R. catesbeiana*, reflecting significant differences in population history or divergent patterns of molecular evolution at mtDNA.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Because of their distribution in a topographically complex and climatically variable environment, widespread temperate species have been the focus of numerous phylogeographic and population genetic studies (reviewed in Taberlet et al., 1998; Brunnsfeld et al., 2001; Soltis et al., 2006). It is generally accepted that the current distribution of mtDNA genetic variation among widespread temperate populations is largely attributable to response to climate change during the Pleistocene through range reductions, population extinctions, and isolation in southern refugia, followed by postglacial range expansions (e.g., Bernatchez and Wilson, 1998; Hewitt, 1999; Zamudio and Savage, 2003). The role of the Pleistocene in the process of speciation has been contentiously debated (Avice et al., 1998; Johnson and Cicero, 2004; Klicka and Zink, 1999). Further, whether co-distributed species responded independently to events in the Pleistocene (Steele and Storfer, 2007), or whether current community structure reflects concordant processes underlying phylogeographic patterns (Lapointe and Rissler,

2005) is a question of interest given the potential role of Pleistocene forces in speciation and the current rate of climatic change.

Comparisons of genealogical histories inferred from mtDNA across a suite of animal species is a prominent field of research (Carstens et al., 2005) despite the problems associated with differences in generation time, rates of divergence, and patterns of sequence evolution potentially confounding interpretation. Because of high mutation rates and matrilineal inheritance in animals, mtDNA has been used extensively for exploring and testing demographic and biogeographic hypotheses (Avice, 2000). However, because the mitochondrion lacks recombination and is inherited as a single unit in most animals, it functions effectively as a single locus. Thus, interpretations based on mtDNA genes are susceptible to error due to lineage sorting and limited precision in estimating population parameters like N_e or divergence times. These shortcomings can be addressed by comparing patterns from multiple independent loci. Although an increasingly common approach, this method is still limited by marker availability and cost (although decreasingly so) for many researchers working on non-model organisms. Further, nuclear sequence markers are often insufficiently variable to resolve recent demographic patterns (e.g., introns that evolve more slowly than mtDNA), or mutate rapidly causing the imprint of historical events such as range fragmentation to be overwritten by homoplasy (e.g., microsatellites). To

* Corresponding author. Address: Wildlife Ecology & Conservation, 110 Newins-Ziegler Hall, P.O. Box 110430, University of Florida, Gainesville, Florida 32611, USA. Fax: +1 352 392 3672.

E-mail addresses: austin1@ufl.edu (J.D. Austin), krz2@cornell.edu (K.R. Zamudio).

partially circumvent this issue one can examine mtDNA genealogical histories across co-distributed species (e.g., Lapointe and Rissler, 2005; Moritz and Faith, 1998; Rowe et al., 2006). These multi-taxa studies can provide insight into how ecological, geological, and historical processes have shaped regional communities. However, because differences in mutation rate and effective population sizes will affect the shape of gene genealogies (but see Gillespie, 2001), using closely related species with similar generation time, vagility, and population ecologies can provide greater insight into the processes that have structured these populations than would comparisons among highly ecologically and evolutionarily divergent species.

In addition to historical demographics, selection can shape DNA polymorphism. Selection potentially impacts population and phylogeographic inference by altering divergence rates and interfering with our understanding of the mode of molecular evolution (Rand and Kahn, 1996; Ho et al., 2005). If undetected, selection can bias the interpretation of population history (Ballard and Whitlock, 2004; Zink, 2005) and potentially confound attempts to examine congruence among species. Given the widespread use of comparative analyses and our efforts to understand how specific habitat, climate, or geographic events shaped genetic variation of species and communities, it is important to understand whether critical assumptions such as neutrality (particularly at protein coding regions, Yang and Bielawski, 2000) are being met, and to further our understanding of what processes (demography and/or selection) are shaping the genetic variation we observe.

Here, we examine intraspecific patterns of divergence in mitochondrial DNA (mtDNA) sequence of two protein-encoding genes in two ecologically similar, co-distributed, and closely related ranid frogs (*Rana clamitans* and *Rana catesbeiana*), that are geographically widespread, and frequently syntopic. The group of seven extant species to which *R. clamitans* and *R. catesbeiana* belong (Dubois, 1992; Hillis and Davis, 1986; Hillis and Wilcoz, 2005) ranges from the Gulf of Mexico to sub-arctic Canada, and species' ranges vary from highly endemic (*R. okaloosae*) to nearly continental distributions (*R. clamitans* and *R. catesbeiana*). *Rana okaloosae* is frequently syntopic with *R. clamitans* in the extreme southern part of the latter's distribution and Austin et al. (2003) found that *R. clamitans* is paraphyletic with respect to *R. okaloosae* mtDNA, suggesting recent divergence of *R. okaloosae* with incomplete lineage sorting, or secondary introgression.

In contrast, the similarity in natural history and ecological preferences of the larger bodied *R. catesbeiana* (SVL 90–150 mm) and *R. clamitans* (SVL 54–100, Conant and Collins, 1998) overlap extensively but differ in frequency across microhabitats (Hecnar and M'Closkey, 1997). For example, *R. clamitans* is found across a range of habitats, from temporary to permanent ponds, whereas *R. catesbeiana* is generally more abundant in permanent water (Werner and McPeck, 1994). The two species have overlapping breeding seasons, similar territorial habits (Ryan, 1980; Wells, 1977) and predatory behaviors (Brooks, 1964; Hamilton, 1948). However, bullfrogs may be important predators of *R. clamitans* (Hecnar and M'Closkey, 1997). Of interest is whether closely related amphibian species sharing similar ecologies will display analogous patterns of divergence and lineage distributions, reflecting similar phylogeographic histories. To address these issues, we examined patterns of mtDNA evolution within the two widespread species, *R. clamitans* and *R. catesbeiana* and ask whether congruent geographic variation exist between the phylogenetically and ecologically similar *R. clamitans* and *R. catesbeiana*. We also ask whether similarities in timing of divergence and past demographics have shaped the two species. Given the possibility of shared mtDNA polymorphism between *R. clamitans* and the endemic *R. okaloosae*, we test specific phylogenetic hypotheses about the relationship of the two species. Finally, because an important assumption behind interpreting

demographic histories based on gene phylogenies is selective neutrality, we explore whether departures from neutrality can attribute to observed patterns.

2. Materials and methods

2.1. Evolutionary units, marker choice, and sequencing procedures

A total of 134 individuals of *Rana clamitans* from 59 localities across its range were included in the initial screen for range-wide mtDNA variation in cytochrome *b* (*cyt b*). We also included in the range-wide *R. clamitans* data set, 11 individuals of *R. okaloosae* from seven localities. This was done as further confirmation of shared polymorphism between *R. clamitans* and *R. okaloosae* (Austin et al., 2003). A *cyt b* analysis of range-wide phylogeographic structure has already been published for *R. catesbeiana* (Austin et al., 2004a). For the purposes of this manuscript we sub-sampled 42 unique *cyt b* haplotypes from Austin et al. (2004a); extending these sequences and added NADH dehydrogenase subunit 2 (ND2) to attempt to improve phylogenetic inference.

Toe clips or tadpole fin clips were collected and stored in 95% ethanol or tissue buffer, or were obtained from tissue collections. Genomic DNA was isolated using DNeasy Tissue Kits (Qiagen). Polymerase chain reaction (PCR) amplification and cycle sequencing was performed on a 1047 bp segment of *cyt b* following methods described in Austin et al. (2003). All 134 *R. clamitans* and 11 *R. okaloosae* were amplified and sequenced for *cyt b* using primers MVZ15-L (Moritz et al., 1992) and MVZar-H (Goebel et al., 1999) to determine range-wide patterns of mtDNA diversity. For both species, we further sequenced a 963 bp segment of ND2 and the flanking tDNA Tryptophan (tDNA^{Trp}) in individuals representing unique *cyt b* haplotypes using primers designed specifically for the *R. catesbeiana* species group (ND2 L4646–5'-ATTGAAGCCTCCA CAAAATA and ND2 H5616–5'-TAAAGGGCCTGAGTTGCATT). PCR methods followed that in Austin et al. (2003), with an annealing temperature of 54 °C.

Amplicons of both gene fragments were used directly for cycle sequencing. Following a clean up with one unit each of exonuclease I and shrimp alkaline phosphatase, cycle sequencing reactions were carried out in 5 µl reactions including 1.5 µl of terminator mix (Applied Biosystems), 0.5 µl 5 × reaction buffer, 0.12 µl primer, and 1–2 µl DNA template and ddH₂O. Sequencing products were brought to 20 µl with ddH₂O and excess dyes and primers were removed using Sephadex purification (Princeton Separations). Cleaned products were electrophoresed on an ABI PRISM® 3100 sequencer (Applied Biosystems). Base calls were checked manually using Sequencher™ vers. 4.5 (Gene Codes); contiguous sequences were created from opposing strands and aligned by eye.

2.2. Intraspecific variation and tests of neutrality

For each gene region, we estimated heterozygosity as the mean number of pairwise differences between all pairs of haplotypes (π) (Tajima, 1983). Because our samples encompass full ranges of both species we estimated π for major haplotype population lineages (see below) as well as range-wide. Net evolutionary divergence among lineages was calculated using the Tamura–Nei method as implemented in Mega vers. 4 (Tamura et al., 2007). This method takes into account differences in rates of substitution between nucleotides and distinguishes between transitional and transversional substitution rates.

To test for deviations from neutral mtDNA evolution we calculated D_{Taj} (Tajima, 1989) using DNAsP vers. 4 (Rozas et al., 2003). Under neutral selection, D_{Taj} is expected to have a mean of zero. We also compared ratios of synonymous and nonsynonymous variation within genes and between *R. clamitans* and *R. catesbeiana*

using the M–K test (McDonald and Kreitman, 1991) and the Neutrality Index (NI) (Rand and Kahn, 1996). The M–K test examines the neutral expectation that the ratio of synonymous and non-synonymous changes should be similar within and between species. Because it is difficult to determine whether one or both species are deviating from neutrality using pairwise M–K tests, we also applied a tree-based method of detecting deviations from neutrality (Creevey and McInerney, 2002). This method was applied to each gene for each species to specifically examine for positive (i.e., adaptive or directional) and negative selection along branches of the phylogenetic trees. First, a rooted NJ tree was estimated from synonymous sequence changes (Zink et al., 2006) and the resulting tree was assumed to represent the true intraspecific phylogeny (Creevey and McInerney, 2002). We implemented the relative rate ratio test (Creevey and McInerney, 2002) by estimating ancestral codon sequences at each branch, and all synonymous and nonsynonymous changes across the tree, using Crann vers. 1.04 (Creevey and McInerney, 2003). The next step involves estimating the number of changes in a descendant clade leading from each internal branch, resulting in four types of substitutions: replacement invariable (RI, new character state is preserved along subsequent lineages); replacement variable (RV, substitutions not preserved in all subsequent lineages); silent invariable (SI, silent changes not observed to have changed again); silent variable (SV, silent changes that have changed at least once). A G-test (or Fisher's exact test when samples are small) was used to compare the ratios SV:RV to SI:RI. As in the M–K test, the ratios are expected to be similar under strict neutrality. In contrast to the M–K test, the relative rate ratio method allows comparison of the relative contribution of RI (indicative of directional selection) to RV (reflecting non-directional selection) (Creevey and McInerney, 2002). In addition, Crann also performs a neutral substitution rate test that identifies instances of negative (purifying) selection. By calculating the total number of replacement and silent sites (Li, 1993) the expected number of replacement and silent sites can be estimated for each clade (Creevey and McInerney, 2002), and excess of deficiency in either type of substitution is compared to chance using a G- or Fisher's exact tests. In the absence of positive selection, significant substitution rate tests represent negative selection (Creevey and McInerney, 2002).

2.3. Phylogenetic analysis

We applied statistical parsimony (Templeton et al., 1992) using TCS 1.21 (Clement et al., 2000) to resolve fine-scale haplotypic relationships among the *R. clamitans* and *R. okaloosae* cyt *b* haplotypes. Cyt *b* haplotypes were chosen because of the larger sample size and to make direct comparisons to cyt *b* in *R. catesbeiana* (Austin et al., 2004a).

Partition homogeneity test, implemented in PAUP, were conducted to confirm homogeneity of signal between cyt *b* and ND2 partitions prior to phylogenetic inference using maximum parsimony (MP) (Camin and Sokal, 1965) for *R. clamitans*/*R. okaloosae* and *R. catesbeiana* data sets using *Rana virgatipes* (Genbank: AY206490, AY083303), *Rana heckscheri* (Genbank: AY206493, AY083299), and either *R. catesbeiana* or *R. clamitans* as outgroup taxa. Heuristic search strategies used stepwise addition with 100 random addition replicates and TBR branch swapping conducted in PAUP* vers. 4 (Swofford, 2002). We evaluated relative support for nodes using 1000 non-parametric bootstrap replicates with 10 random additions per replicate. Due to the large number of *R. clamitans*/*R. okaloosae* haplotypes with low divergences, we used a pruned data set for maximum parsimony analyses, omitting rare haplotypes that were interiorly located (i.e., non-tip haplotypes) and that differed by other combined cyt *b*-ND2 haplotypes only 1–2 bp (only relatively non-informative haplotypes).

We examined various models of DNA evolution for each gene region separately and combined using Modeltest vers. 3.7 (Posada and Crandall, 1998). Bayesian analysis was implemented using MrBayes vers.3.1.2 (Huelsenbeck and Ronquist, 2001) using the best fit model parameters and default priors. We performed two concurrent Bayesian searches, each of 5.0×10^6 generations consisting of four chains with a heating parameter of 0.5 and all other parameters set to default. Standard deviations of split frequencies were examined to determine run convergence, and the plot of log likelihoods versus generation was examined to determine stationarity. Based on these results, we discarded the first 2.5×10^5 generations as burn in. We considered posterior probability values to be 'significant' at 0.95 or greater (Leache and Reeder, 2002).

2.4. Parametric test of lineage divergence

Due to the limited support for basal lineages using traditional phylogenetic methods we applied parametric bootstrapping to test the monophyly of diagnosed lineages within species and to explore the strength of phylogenetic hypotheses derived from MP and Bayesian algorithms. We compared the parsimony tree score differences obtained under a constraint and unconstrained topology, with the same differences obtained from 100 simulated datasets (Huelsenbeck et al., 1996). Simulated sequences were subsequently constructed using SG Runner vers. 1.5.3 (T.P. Wilcox, unpublished) and Seqgen vers. 1.3.2 (Rambaut and Grassly, 1997). Sequences were simulated using the same model of DNA evolution estimated from the observed sequence data, modeled over the tree topology and branch lengths from the constrained tree. Each simulated data set was subject to constrained and unconstrained heuristic searches in PAUP*. For each replicate data set, tree scores were compared between constrained and unconstrained searches, producing a null distribution against which to test the significance of our original tree score differences. An empirical difference between constrained and unconstrained tree scores that falls outside the 95% tail of the null distribution means we can reject the null hypothesis.

2.5. Lineage divergence model

Divergence among lineages within species was estimated two ways, both methods were conducted on the entire cyt *b* gene datasets because of their larger sample sizes (number of individuals and numbers of haplotypes) than combined cyt *b*-ND2 datasets (see Austin et al., 2004a for information on *R. catesbeiana* cyt *b*). First, we calculated net TN (Tamura and Nei, 1993) among lineage divergence \pm standard error using MEGA4 (Tamura et al., 2007). Net among clade divergence produces an unbiased estimate assuming lineages are reciprocally monophyletic and that the ancestral population was equal to the average of the descendent lineages (Arbogast et al., 2002).

To address whether divergence between major mtDNA lineages of *R. clamitans* and *R. catesbeiana* occurred at similar times, and with similar demographic consequences, we used the isolation with migration model described by Nielsen and Wakeley (2001) implemented in the program IMA (Hey and Nielsen, 2007). IMA estimates posterior probabilities for demographic model parameters; in this case, we estimated the ancestral effective population size, θ_A (prior to lineage divergence), and the effective population sizes of the two daughter lineages, θ_1 and θ_2 , reflected in our phylogeographic analyses. Our interest was in estimating relative θ and t (time since divergence scaled by μ) for the two major lineages in each species (see below), assuming that within each species these lineages had diverged in allopatry. IMA also can examine the influence of gene flow; however, we ignored the migration parameter (i.e., set migration to 0) because we are not dealing with

discrete geographic populations and because we are only looking at matrilineal history. Because we are only examining a single locus, reducing the number of parameters increases the ability of the Markov chain Monte Carlo (MCMC) to explore parameter space.

We implemented the HKY mutation model and an appropriate mtDNA inheritance scalar of 0.25 (i.e., the expected effective population size relative to an autosome). To optimize number of chains, heating model, and parameter bounds for our long runs, we ran a series of pilot metropolis coupled Monte Carlo Markov chains to refine our analyses. Our final run consisted of 5.0×10^5 burnin steps followed by a minimum of 3.0×10^7 steps, using a linear heating scheme consisting of three heated chains. Stationarity was assessed by examining plots of parameter trends and by ensuring effective sample sizes (ESS) were greater than 50. We omitted comparisons between Louisiana haplotypes ($N = 2$) and widespread or Coastal Plain lineages due to the small sample size of the former. Preliminary IMA runs suggested that including comparisons of Louisiana would not likely achieve reliable estimates.

Peaks of posterior distributions were used as maximum-likelihood estimates of the parameters of interest (Nielsen and Wakeley, 2001). We used the 90% highest posterior density (HPD) interval as our estimates of credibility intervals. Posterior distributions are only relatively comparable due to their scaling by μ , the mutation rate per gene. Because we wish to compare patterns between species using different lengths of *cyt b* we incorporated a range of μ estimated for *cyt b* and similar genes for poikilothermic organisms. In the *R. catesbeiana* species group, *cyt b* and ND2 evolve at similar rates (Austin et al., 2003), therefore, we incorporate studies that estimate divergence based on either gene. We used estimates of *cyt b* divergence ranging from 0.69% to 2.4% per million years (Macey et al., 1998; Plötner et al., 2001), as these encompass a range of estimated rates for poikilotherms (Martin and Palumbi, 1993; Tan and Wake, 1995; Caccone et al., 1997; Mueller, 2006). These rates translate into mutation rate/year/gene region of 5.66×10^{-6} – 1.70×10^{-5} for *R. clamitans* (821 bp of *cyt b*) and 2.78×10^{-6} – 9.82×10^{-6} for *R. catesbeiana* (409 bp *cyt b*). These values were

used as upper and lower bounds for estimates of N_e and divergence time. Because the coalescent model defines N_e as being proportional to the inverse of the generational coalescent rate (Hey and Nielsen, 2004) we adopted an average generation time of three years for both species (Martof, 1956; Durham and Bennett, 1963; Schroeder and Baskett, 1968).

3. Results

3.1. Genetic variation and lineage identification

We obtained 821 bp of aligned *cyt b* sequence from 134 *R. clamitans* and 11 *R. okaloosae* representing 59 localities (Fig. 1). Protein coding regions were translated into amino acids, and a bias against guanine on the light strand (Table 1) support the authenticity of mtDNA (Brown, 1985). We recovered 57 unique *cyt b* haplotypes [Genbank: DQ826040–DQ826043, DQ792651–DQ792704] with pairwise sequence divergence (HKY) varying from 0.1% to 3.4%. A subset of individuals ($N = 64$) representing all unique *cyt b* haplotypes were sequenced for an additional 796 bp of ND2 and 49 bp of tRNA^{Trp}. From the 64 individuals sequenced for ND2 we recovered 33 unique haplotypes with HKY pairwise sequence divergence varying from 0.01% to 3.1% [Genbank: DQ792651–DQ792704]. For *R. catesbeiana* we extended 41 previously published *cyt b* haplotypes (Austin et al., 2004a) for a total of 884 aligned bp (varying from 0.01% to 2.6% HKY distance), plus 893 bp of ND2 + tRNA^{Trp}, varying from 0.01% to 1.6% pairwise HKY sequence divergence [Genbank: EF122794–EF122834]. Partition homogeneity tests were not significant for either *R. clamitans* or *R. catesbeiana* ($P > 0.05$).

3.2. Tests for selection

Negative D_{Taj} results from $\theta > \pi$, and in our tests, this occurred more than expected by chance in all species and lineages (sign test, $P < 0.001$, Table 2). Nonetheless, the relationship between θ and π was close to the expectation under neutrality for *R. catesbeiana* in

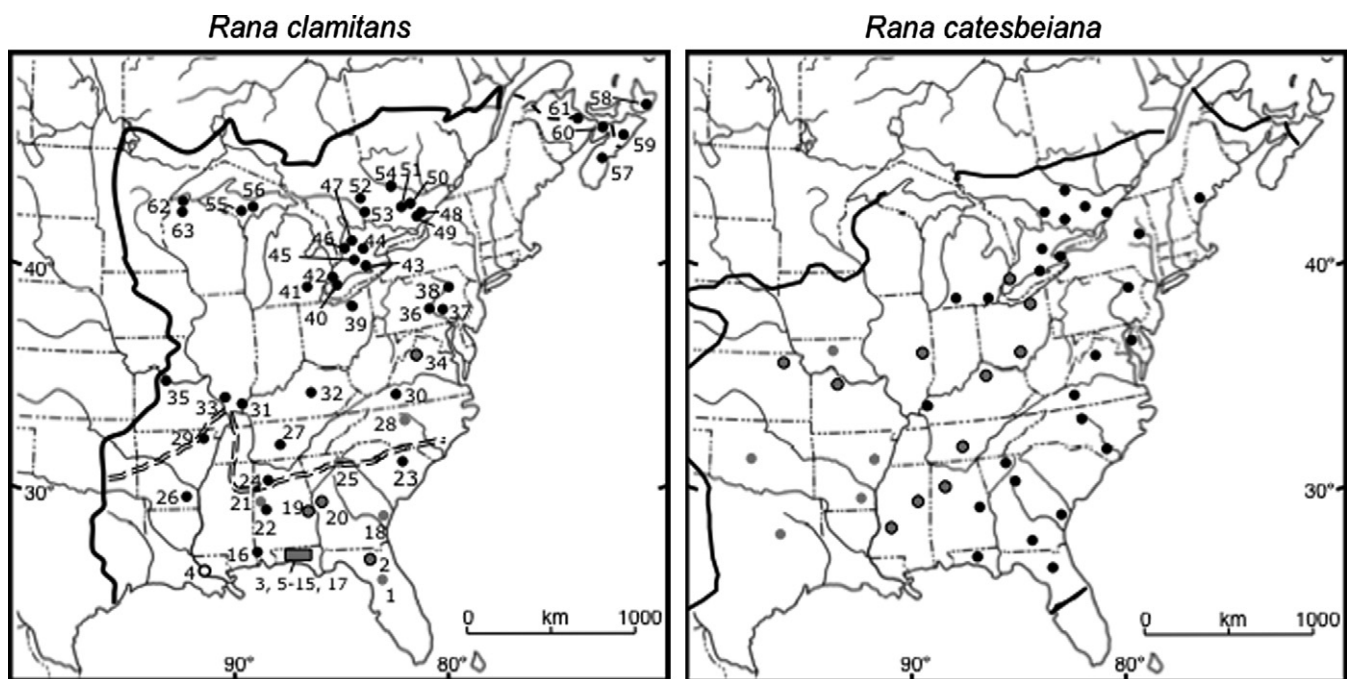


Fig. 1. Range distributions (solid lines) and sampling localities for *R. clamitans*/*R. okaloosae* and *R. catesbeiana* (redrawn from Austin et al. (2004a)). Localities with haplotypes belonging to each of two major lineages are indicated by circles: black (*R. clamitans*, widespread; *R. catesbeiana*, eastern) gray (*R. clamitans*, Appalachian-Coastal Plain; *R. catesbeiana*, western), or Grey with black outline (both lineages present). Numbers in *R. clamitans* refer to locations in Supplementary file 1. Location 4 (Louisiana) contained divergent haplotypes. The rectangle encompasses the range of *R. okaloosae* and includes 13 sampling sites. Dashed lines represent the approximate boundary (Mecham 1954) separating putative subspecies *R. c. melanota* (north) and *R. c. clamitans* (south).

Table 1
Sequence data and model parameter results

	ND2	tRNA-Trp	cyt <i>b</i>	Combined
<i>L. clamitans</i>				
Aligned bp	796 ([†] 812–1608)	49 ([†] 1609–1658)	821 (^{††} 95–914)	1666
Base frequencies (A:C:G:T)	0.248:0.300:0.135:0.316		0.248:0.300:0.135:0.316	0.2687:0.2965:0.1219:0.3129
Prop invariable sites	—		0.508	0.524
Gamma shape parameter	0.182		0.721	0.7525
Model selected	GTR+G		HKY+I+G	TIM+I+G
<i>L. catesbeiana</i>				
Aligned bp	821 ([†] 786–1608)	72 ([†] 1609–1680)	884 (^{††} 60–944)	1777
Base frequencies (A:C:G:T)	0.280:0.277:0.124:0.319		0.239:0.313:0.136:0.312	0.2635:0.2953:0.1291:0.3121
Prop invariable sites	0.8558		—	0.433
Gamma shape parameter	Equal rates		0.421	0.8595
Model selected	TIM+I		HKY+G	TrN+I+G

[†] Homologous bp position relative to GenBank No. AF314016.

^{††} Homologous bp position relative to GenBank No. AF205089.

Table 2
Nucleotide diversity and estimate of neutrality in ND2 and cyt *b*

Gene	Species	<i>S</i>	π	D_{Taj}	$d_N:d_S$
ND2	<i>R. clamitans/okaloosae</i>				
	CPA lineage	27	0.006	−1.24	6:22
	Widespread lineage	19	0.002	−2.16 [*]	3:19
	Louisiana	4	0.005	—	0:04
	Combined	52	0.007	−1.87 [*]	9:32
	<i>R. catesbeiana</i>				
	East lineage	25	0.005	−1.67	6:19
cyt <i>b</i>	<i>R. clamitans/okaloosae</i>				
	CPA lineage	39	0.007	−1.57	6:33
	Widespread lineage	35	0.005	−1.90 [*]	5:31
	Louisiana	1	0.001	—	0:01
	Combined	80	0.012	−1.63	11:57
	<i>R. catesbeiana</i>				
	East lineage	38	0.007	−1.53	12:26
Combined	<i>R. clamitans/okaloosae</i>				
	CPA lineage	69	0.007	−1.57	8:40
	Widespread lineage	48	0.003	−1.96 [*]	14:56
	Louisiana	5	0.003	—	0:05
	Combined	134	0.009	−1.78	27:143
	<i>R. catesbeiana</i>				
	East lineage	80	0.01	−1.01	25:56
West lineage	60	0.011	−0.33	13:46	
	Combined	102	0.01	−1.16	32:71

S, number of segregating sites.

π , mean number of pairwise differences between all haplotypes.

D_{Taj} , Tajima (1989) test for departure from neutrality ($P < 0.05$).

$d_N:d_S$, ratio of non-synonymous to synonymous sites.

both gene fragments, but not for *R. clamitans* where we detected significantly low D_{Taj} values for the widespread lineage in both genes and across all haplotypes for ND2.

Results from M–K tests on individual and combined gene fragments in *R. clamitans* widespread, CPA and Louisiana lineages (below), and *R. catesbeiana* east and west lineages were not significant (Table 3). In contrast, interspecific M–K tests suggest deviations from neutrality in cyt *b*, but not ND2. The combined gene M–K test was also significant (not shown, G test = 5.37, $P = 0.01$).

Tree-based estimates of selection show instances of positive selection; most frequently at the branches separating ingroup lineages from outgroups in both species for cyt *b* (Fig. 2). Non-directional positive selection was also inferred (indicated by replacement invariable (RI) substitutions < replacement variable (RV) substitutions, Supplementary files 1 and 2), at the basal split

Table 3
McDonald–Kreitman tests for individual protein coding genes

			Widespread		CPA		
			Fixed	Variable	Fixed	Variable	
<i>R. clamitans</i>	ND2	CPA	Synonymous	0	32		
			Nonsynonymous	0	9		
			Synonymous	8	18	8	24
			Nonsynonymous	4	3	5	6
cyt <i>b</i>	CPA	Synonymous	4	57			
		Nonsynonymous	0	11			
		Synonymous	14	32	15	34	
		Nonsynonymous	0	5	0	6	
<i>R. catesbeiana</i>	ND2	East	West				
			Fixed		Variable		
			Synonymous	3	25		
			Nonsynonymous	0	10		
cyt <i>b</i>	East	Synonymous	5	39			
		Nonsynonymous	0	20			
<i>R. clamitans</i> vs. <i>R. catesbeiana</i>	ND2	<i>R. clamitans</i>	Fixed				
			Variable		Variable		
			Synonymous	29	62		
			Nonsynonymous	9	22		
cyt <i>b</i>	<i>R. catesbeiana</i>	Synonymous	38	112			
		Nonsynonymous	0	31		$P < 0.001$	

Significant values in italics.

between Louisiana haplotypes and the rest of the *R. clamitans* lineages, and in *R. catesbeiana* at the branch leading to the split between east and west lineages, and basally within the west lineage. One instance of positive directional selection was detected in *R. clamitans* at the ND2 locus along the branch leading to the Louisiana haplotypes. Significant deficits of replacement changes, indicative of negative selection, were detected across both genes in both species, most of which occurred basally in the trees.

3.3. Phylogenetic relationships

All but two *R. clamitans* cyt *b* haplotypes connected in a statistical parsimony network within the 95% limits of parsimony (12 steps). Two haplotypes (36 and 37) from Louisiana differed from all others by 13 and 15 steps, respectively (Fig. 3). All 11 *R. okaloosae* haplotypes are shared by or nested within *R. clamitans* (Fig. 3).

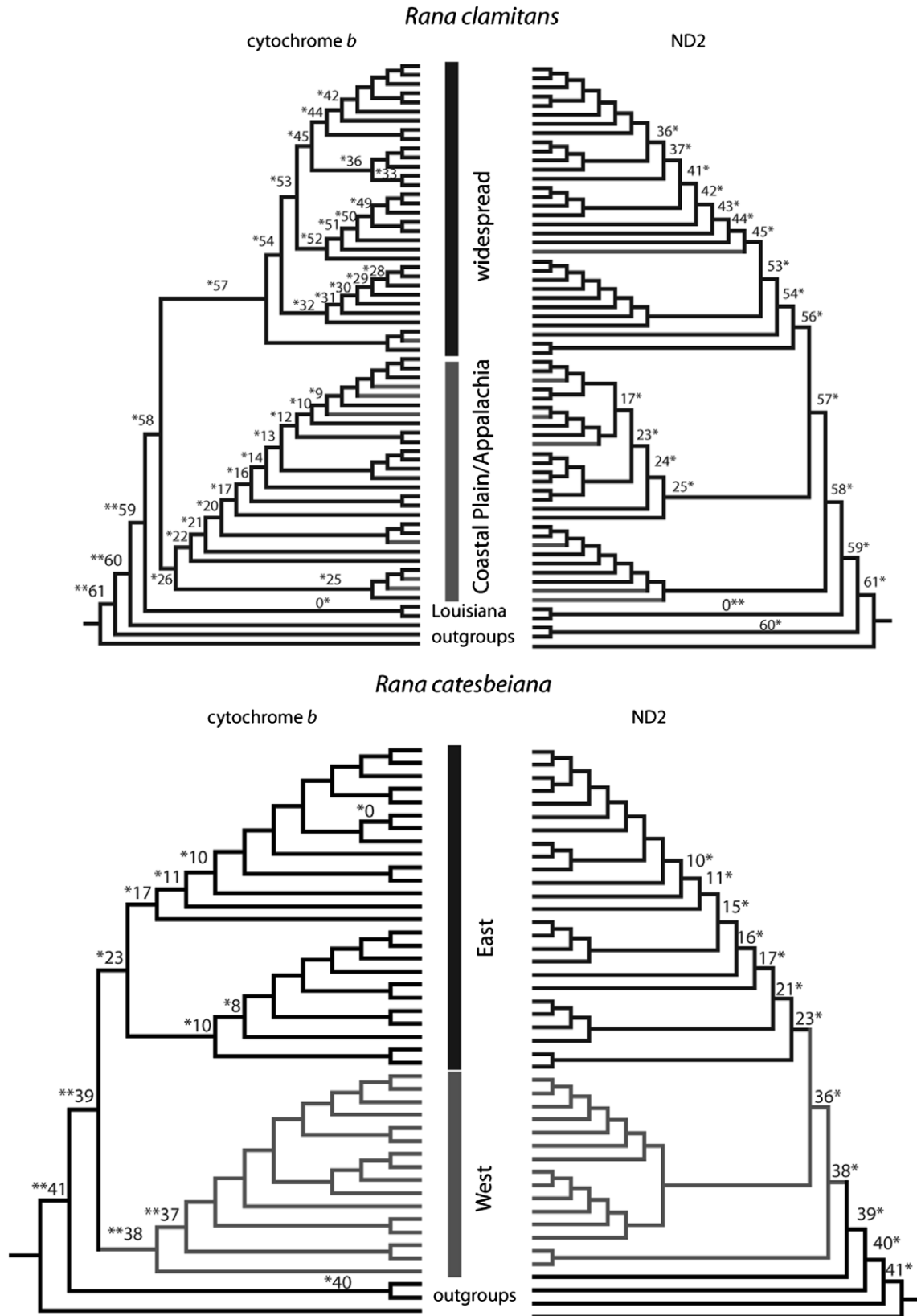


Fig. 2. NJ trees (distances estimated from synonymous changes only) for *cyt b* and ND2 genes. Double asterisks indicate significant differences in the ratios of RI/RV to SI/SV indicating positive selection. Single asterisks indicate significant deficiency in the number of replacement substitutions expected under neutrality. Only numbers for significant branches are shown to reduce clutter (see Supplementary files 1 and 2). Gray branches in *R. clamitans* trees indicate position of *R. okaloosae* haplotypes.

Therefore, for brevity, we refer to all of these as *R. clamitans* haplotypes, although this does not reflect our opinion on the species

status of *R. okaloosae* or *R. clamitans*. Discussion of the taxonomic status of *R. clamitans* is beyond the scope of this paper.

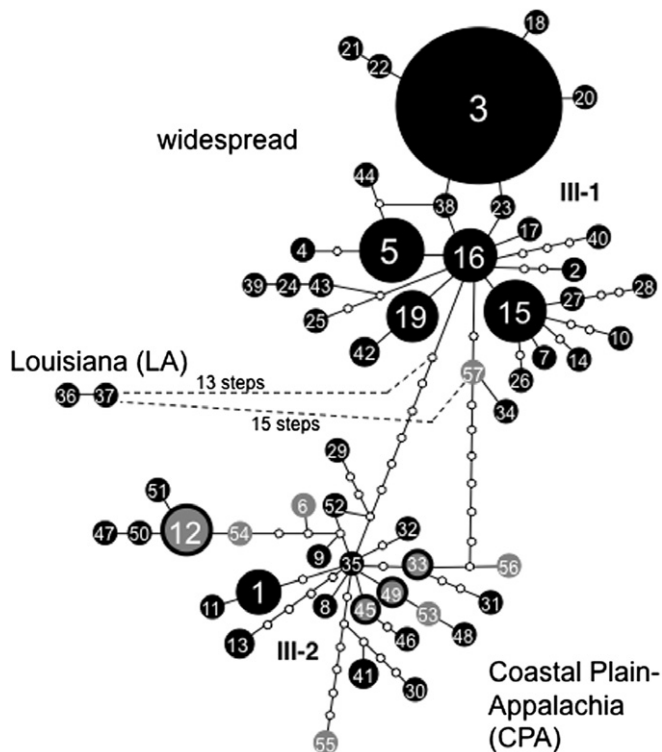


Fig. 3. Statistical parsimony network of 57 *Rana clamitans* and *R. okaloosae* haplotypes derived from 821 bp of *cyt b* sequence data. *R. clamitans* haplotypes are black, *R. okaloosae* haplotypes are gray, and shared haplotypes are indicated by both black and gray.

Within the main *R. clamitans* haplotype network we identified two primary mtDNA lineages separated by a minimum of seven bp differences. Geographically, the distribution of these two lineages represent a Coastal Plain–eastern Appalachian (CPA) lineage and a widespread haplotype group (widespread) that overlapped with the CPA lineage (Fig. 1). *Rana okaloosae* haplotypes were primarily part of the CPA haplotype group, with the exception of haplotype 57 that was nested in the widespread lineage, separated from the closest *R. okaloosae* haplotype by 7 mutational steps. Four *cyt b* haplotypes were shared between *R. clamitans* and *R. okaloosae* (Fig. 3). In comparison to the pattern observed in *R. clamitans*, the previously published *R. catesbeiana* network also consisted of two main clades (East and West) separated by a minimum of five mutational steps (Fig. 3 in Austin et al., 2004a). Net sequence divergence based on combined gene partitions was 0.53% between widespread and CPA lineages, and 2.1% and 1.9% between the Louisiana (LA) and CPA, and LA and widespread lineages, respectively. Net divergence was 0.63% for *R. catesbeiana* east and west clades.

Maximum parsimony analysis of the *R. clamitans* haplotypes based on 1666 bp of combined *cyt b* and ND2 resulted in 2.27×10^4 equally parsimonious trees of 514 steps. The MP consensus tree reflected the parsimony network in that the node separating the LA from all other clades and the nodes leading to the CPA and widespread lineages had high to moderate support (100–72%, respectively, Fig. 4A). The node grouping the CPA lineage had moderate support (75%), with low to good support for various branches within this clade (69–100%), including clades that contained most *R. okaloosae* haplotypes. In contrast, the widespread lineage was characterized by limited structure and low bootstrap support (<65). The *R. catesbeiana* MP tree based on a combined 1777 bp resulted in 94 equally parsimonious trees 518 steps in length. There was moderate to strong support (71–94%) for eastern and western clades (Fig. 4B).

Modeltest predicted various best-fit models of DNA evolution among gene partitions and combined gene sequences for both *R. clamitans* and *R. catesbeiana* (Table 1). However, tree topologies incorporating different models of DNA evolution did not differ substantially (only small variation in terminal branch topologies); we therefore limit our results and discussion to the combined model results. The phylogram resulting from 15,000 samples from two concurrent *R. clamitans* Bayesian runs (discarding 2500 samples from each run as burn in) only weakly resembled the MP topology in that the Louisiana haplotypes were placed together with high support but were not differentiated basally from widespread nor Coastal Plain lineages (not shown). The node leading to the CPA lineage was well supported (posterior probability ≥ 0.95), with significant posterior support defining nodes leading to shared *clamitans/okaloosae* haplotypes. Support for the widespread clade was less than 0.95. Bayesian results from *R. catesbeiana* produced a phylogram with strong support for the eastern clade; however, western haplotypes identified in the MP analysis formed a basal polytomy in the Bayesian phylogram similar to that produced by *cyt b* alone (Austin et al., 2004a). Overall, despite weak support for some clades in either or both MP and Bayesian analyses, the results from the parsimony network analysis suggest that lineages in each species had been reproductively isolated for some period in the past. Given the low level of divergence and frequency of common haplotypes in each species, a non-bifurcating tree approach (i.e., network) is likely to be the most appropriate for examining population history within species with shallow evolutionary histories (Posada and Crandall, 2001).

3.4. Hypothesis testing of various tree topologies

Due to the varying phylogenetic support for the major clades discussed above, we used parametric bootstrapping to further estimate support for the monophyly of particular geographic clades within each data set. We tested the significance of four constrained hypothetical topologies with: (1) reciprocal monophyly of the widespread and CPA lineages (omitting the 2 LA haplotypes), (2) monophyly of *R. okaloosae* and *R. clamitans*, (3) monophyly of *R. okaloosae*, paraphyly of *R. clamitans*, (4) monophyly of the east and west *R. catesbeiana* clades. We could not reject the hypotheses of monophyly for the major lineages within either *R. clamitans* or *R. catesbeiana*. However, both topologies constraining *R. okaloosae* as a monophyletic lineage were statistically rejected ($P < 0.001$, Fig. 5).

3.5. Coalescent estimates of population structure

The ESS values from our final IMA runs ranged from 60 to greater than 10,000. Although some parameters (usually t) had relatively low ESS values (60–100), examinations of parameter trend plots suggested good parameter space exploration and posterior estimates generated during our preliminary and final runs were very similar. Fig. 6 shows the posterior probability distributions of the estimates of N_e values for the two lineages and their common ancestor within each of *R. clamitans* and *R. catesbeiana*. Results from both species indicate population growth since lineage splitting because each lineage has larger N_e than the ancestral N_e . In both species the geographically more widespread lineage ('widespread' and 'east' lineages) have larger N_e than the more restricted lineages (Fig. 6). The *R. clamitans* widespread θ likelihood estimate is less resolved, having a broad distribution relative to the CPA lineage and to either of the contemporary *R. catesbeiana* estimates. The *R. catesbeiana* east lineage had a larger θ value than the west, although with some overlap of their HPD. The ancestral population parameter θ_A shows a relatively narrow distribution and distinct peak in *R. clamitans* relative to *R. catesbeiana*. In *R. catesbeiana*,

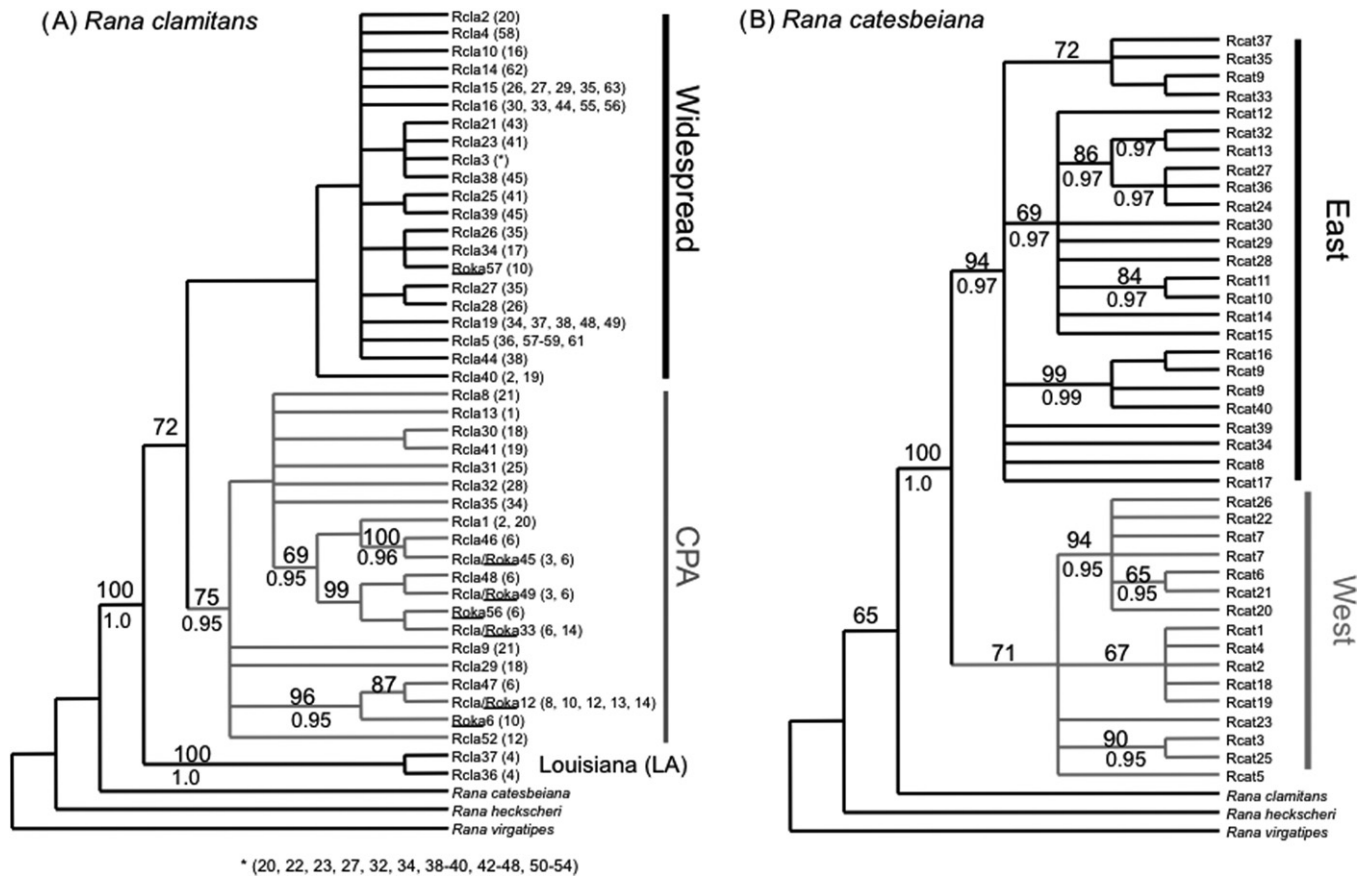


Fig. 4. Maximum parsimony cladograms with non-parametric bootstrap support (1000 replicates) based on combined ND2-tDNA^{Trp} and *cyt b* fragments: (A) *Rana clamitans*/*R. okaloosae*, (B) *R. catesbeiana* haplotypes. Identifiers for *R. clamitans* are based on *cyt b* haplotypes identified in Fig. 2, and locations in parentheses correspond to map in Fig. 1. Haplotypes for *R. catesbeiana* correspond to Fig. 3 of Austin et al. (2004a). Numbers above branches are nonparametric bootstrap support greater than or equal to 65. Bayesian posterior support (≥ 0.95) for the corresponding is indicated below branches.

the curve for θ_A is broad and overlaps with high probability with zero. Assuming three year generation times on average and a mutation rate of 0.69% per million years, effective matrilineal population size of the *R. clamitans* ancestral population [470,000 (HPD 174,000–1,200,000)] increased 4-fold in the CPA [1,900,000 (HPD 1,300,000–2,700,000)] and eight times in the widespread lineage [4.0×10^6 (HPD 2.5×10^6 – 6.4×10^6)]. In contrast, *R. catesbeiana* lineages ranged from 3.8×10^6 (2.5×10^6 – 5.5×10^6) to 1.9×10^6 (1.0×10^6 – 3.3×10^6) for the slowest mutation rate. Although a distinct peak was observed for the ancestral *R. catesbeiana* population, the distribution was very broad and a large amount of the likelihood distribution overlapped with zero (Fig. 6B), making comparisons to the descendent lineages difficult. The effect of using the faster mutation rate (2.4% pmy) was to reduce estimates of N_e by more than half (Fig. 6A and B).

Estimates of time since divergence between lineages of *R. clamitans* based on coalescent estimates of t , reveal a sharp peak marginal likelihood in *R. clamitans* that translates into a period ranging from the onset of the last Wisconsin glaciation (115,000 years ago), assuming a fast molecular clock (2.4%), to a divergence estimate of approximately 400,000 years ago with a slower poikilothermic clock of 0.68% per million years. In contrast, the distribution of marginal likelihood values of t in *R. catesbeiana* was broad but gives older estimates of divergence than in *R. clamitans* (Fig. 6C) ranging from approximately 380,000 years ago to 1.3 Mya.

Divergence times based on net divergence (Table 4) produced older estimates compared to coalescent based estimates produced by IMA. However, these estimates were consistent with the knowl-

edge that net divergence estimates time to most recent common ancestor, while t , estimated in IMA, reflects time since population splitting.

4. Discussion

The main objective of this study was to compare the mtDNA population history between closely related anuran species with widely overlapping distributions and similar ecologies. Assuming no interspecific differences in patterns of selection and evolutionary rates, these species could have displayed similar gross-scale population genetic structure, reflecting similar responses to past bioclimatic changes and physiographic features. This congruence hypothesis assumes that evolutionary rates are approximate and primarily affected by biological systems (Thorne and Kishino, 2005). It follows that by comparing closely related species with similar ecologies we can approach the problem of comparing evolution histories of specific genes with more confidence than by comparing dissimilar species.

Our comparative mtDNA analysis of *R. clamitans* and *R. catesbeiana* sampled from across eastern North America revealed two main widespread and partially overlapping lineages in each species. However, there appears to be little geographic concordance between *R. clamitans* and *R. catesbeiana* (Fig. 1) that would suggest that physiographic features could explain the distribution of mtDNA lineages. Both species are likely good dispersers (Austin et al., 2004b; Schroeder, 1976; Willis et al., 1956) and therefore, features that have been used to explain the geographic structuring

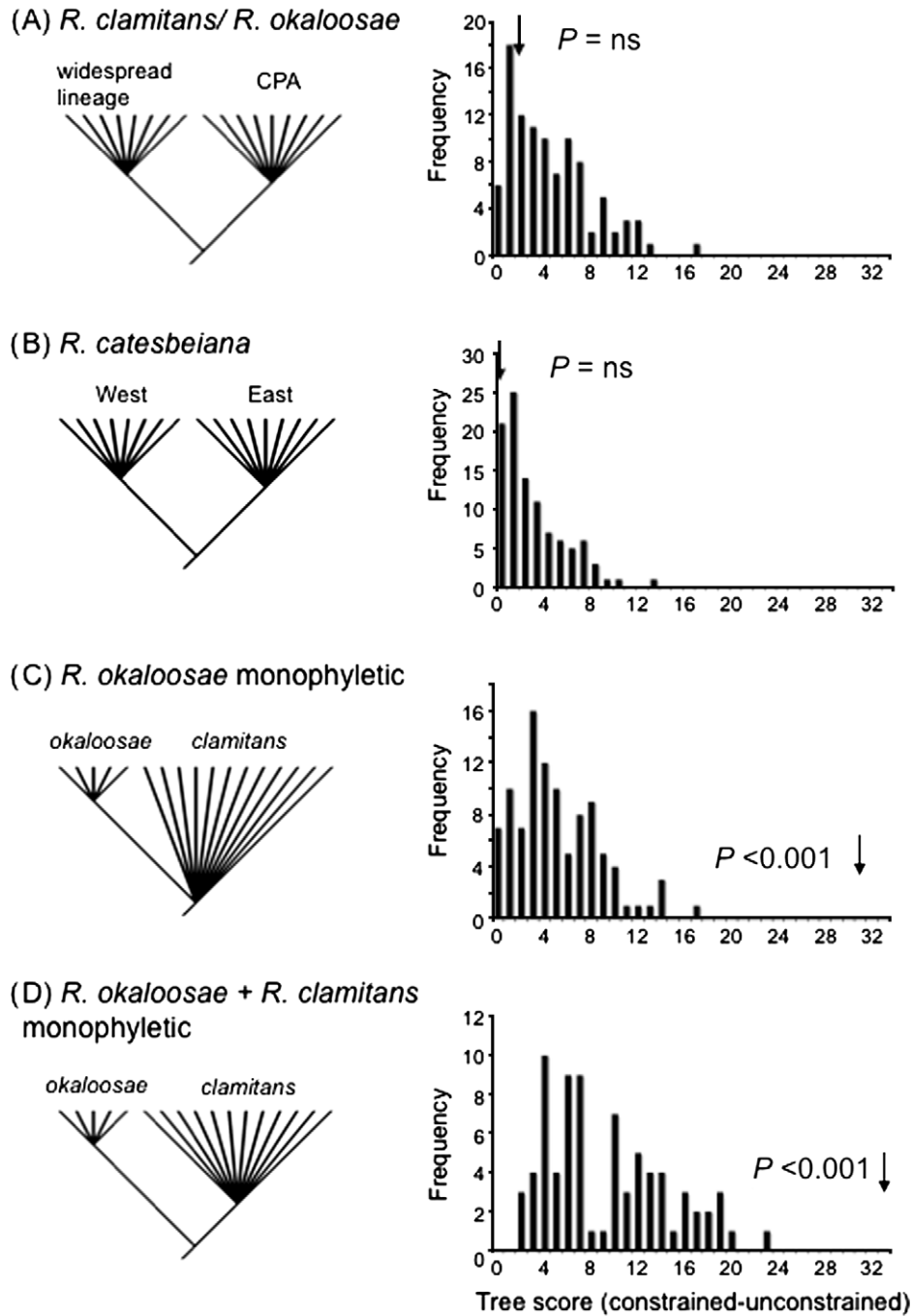


Fig. 5. Null distributions of the parametric bootstrap tests designed to test support for phylogenetic hypotheses derived from network, MP and Bayesian analyses (A and B), and to test alternative hypotheses regarding the relationship among *R. okaloosae* and *R. clamitans* (C and D). (A) Monophyly of widespread and CPA lineages in *R. clamitans*/*R. okaloosae*. (B) Monophyly of east and west lineages of *R. catesbeiana*. (C) *R. okaloosae* constrained to be monophyletic. (D) Both *R. okaloosae* and *R. clamitans* constrained to be monophyletic.

of widespread but relatively low vagility species are likely to have less impact on genetic structuring.

Austin et al. (2004a) argued that the west lineage of *R. catesbeiana* became isolated west of the Mississippi River during the Pleistocene. The presence of highly differentiated *R. clamitans* haplotypes in Louisiana corresponds to this pattern of isolation in a southwestern refugial area; unfortunately our sampling in this portion of the range limits our ability to examine this pattern in any detail. However, the presence of an ancestral region (interpreted as being a refugial area) in the coastal plain area west of the Mississippi has been inferred to be an important area of diver-

sification for a variety of widespread taxa (e.g., Al-Rabab'ah and Williams, 2002; Howes et al., 2006; Zamudio and Savage, 2003). With the exception of the Louisiana haplotypes, there is little to support that neither the CPA nor the widespread lineage of *R. clamitans* was isolated west of the Mississippi. The widespread lineage is found west of the Mississippi and across the remainder of the *R. clamitans* range, including the distribution of the CPA, which is restricted to east of the Mississippi and south and east of the fall line (Appalachian Mountains). This relative distribution, together with the smaller population size inferred from IMA and the lack of evidence from D_{Taj} (Table 2) suggests that the CPA has undergone rel-

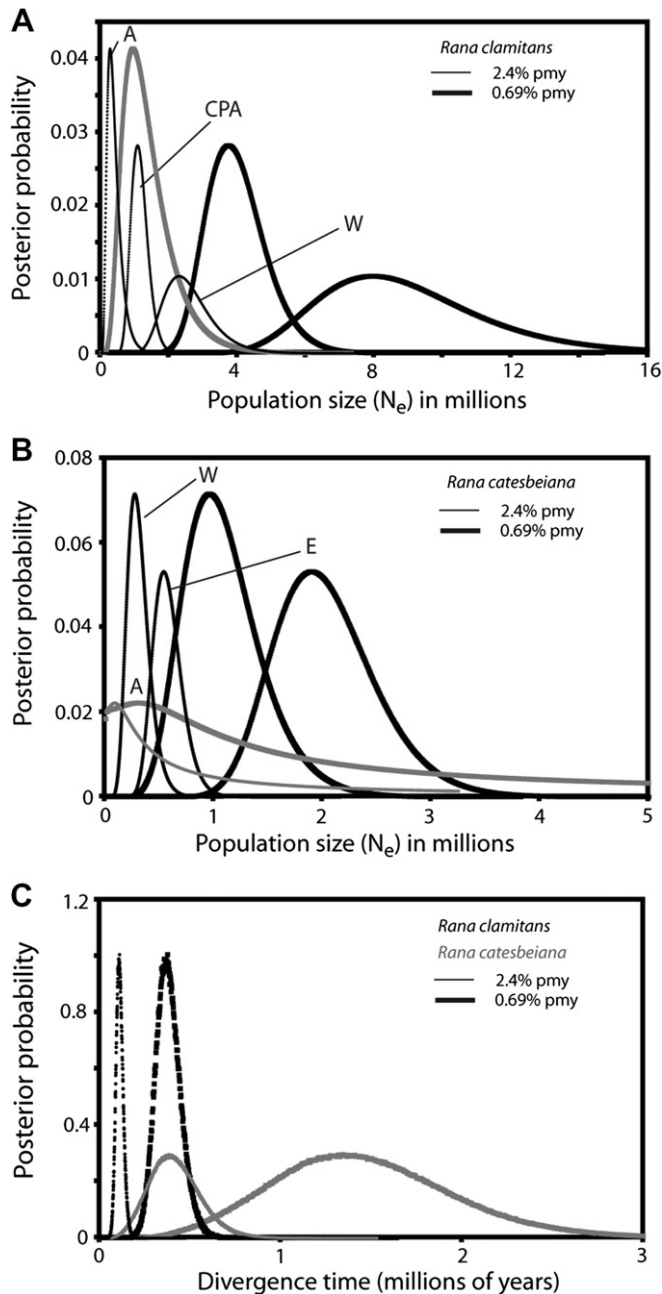


Fig. 6. Marginal posterior probability distributions for demographic estimates of population size and time since population divergence. (A) *Rana clamitans*: Estimates of N_e for the widespread (W), CPA, and ancestral (A, gray lines) lineages at two mutation rate scalars. Thin lines represent 2.4% per million years; thick lines are 0.69% per million years. (B) *Rana catesbeiana*: Estimates of N_e for the east (E), west (W), and ancestral (A, gray lines) lineages at two mutation rate scalars. Thin lines represent 2.4% per million years; thick lines are 0.69% per million years. (C) Time since lineage divergence in *Rana clamitans* (black) and *Rana catesbeiana* (gray). Thin lines represent 2.4% per million years; thick lines are 0.69% per million years.

Table 4
Time of divergence among lineages based on net among lineage divergence

Lineage	Lineage comparison	MYA (\pm SE) at 0.6%	MYA (\pm SE) at 0.8%
<i>R. clamitans</i>	Widespread vs. CPA	1.61 (0.49)	1.21 (0.37)
	Widespread vs. LA	3.42 (0.84)	2.56 (0.63)
	CPA vs. LA	4.04 (0.89)	3.03 (0.67)
<i>R. catesbeiana</i>	East vs. west	2.48 (0.93)	1.86 (0.70)

actively less population growth or expansion over the Pleistocene. In contrast, the widespread lineage shows patterns that typically

indicate range expansion, namely limited resolution among haplotypes (Fig. 4), distribution well into previously glaciated regions, and significant negative D_{Taj} values. The distribution of the widespread lineage suggests either expansion into the Coastal Plain from a separate refugial area, or expansion out of the area while the CPA has remained relatively restricted geographically.

Temporal analyses using IMA suggest that the differences in geographic patterns may be related to inter-specific differences in timing of major lineage bottlenecks or isolation. If we assume similarity in divergence rates, mtDNA lineages from these species coalesce over different time scales, with *R. catesbeiana* diverging earlier in the Pleistocene than *R. clamitans*. It is expected that the signature of positive selection would have minimal impact on divergence estimates because our tree-based estimates of selection is most prevalent at the inter-specific boundary. Examinations of the effect of selection on the coalescent (Neuhauser and Krone, 1997; Barton and Etheridge, 2004) suggest that there should be a negligible impact of weak selection on the topology of gene genealogies. However, strong selection can have a substantial effect on the length of a genealogy particularly deep in the tree, and thus on the inference of timing of diversification. Basally, the relatively small numbers of lineages means that fluctuation in allele frequency, that mitigates the effect on selection, has minimal influence (Barton and Etheridge, 2004). The relative greater frequency of positive selection detected basally in *R. catesbeiana* (Fig. 2) may be influencing the estimate of t or θ_A in IMA (below).

Austin et al. (2004a) inferred allopatric differentiation of the two main *R. catesbeiana* lineages based on the greater than average number of bp differences and the relatively restricted geographic distributions of basal haplotypes within both east and west lineages. However, phylogenetic reconstructions did not support the reciprocal monophyly of these clades, nor is monophyly well resolved within *R. clamitans* using traditional tree-based methods, although there is a larger than average number of mutations separating the two clades (Fig. 3). Our parametric bootstrap test, that simulates sequence evolution within these two taxa, further supports the null hypothesis of a monophyletic origin of the sequences in each of the main clades in either species despite the poor resolution provided by nonparametric bootstrap or Bayesian methods. The extensive overlap in the distribution of the CPA and widespread lineages suggest that the divergence observed between the two is not merely due to insufficient sampling causing isolation by distance to be mistaken for allopatric divergence (Bridle et al., 2004). Further, the presence of highly divergent haplotypes in Louisiana suggests that southwest portion of the range of *R. clamitans* may harbor a basal lineage that is geographically similar to the ancestral region inferred for the *R. catesbeiana* west lineage. Formal testing of this hypothesis awaits larger sample sizes from the Louisiana lineage.

Although phenotypic variation has been used to distinguish northern putative subspecies (*R. clamitans melanota*) from a southern subspecies (*R. clamitans clamitans*, Mecham, 1954; see Fig. 1A), these distinctions are not reflected by the distribution of the mtDNA lineages identified here. Numerous recent phylogenetic studies of widespread North American taxa have documented a lack of congruence between mtDNA and subspecies designations based on external phenotypic polymorphisms (e.g., Burbrink et al., 2000; Clark et al., 2003; Moriarty and Cannatella, 2004).

The divergence in advertisement call structure and external phenotypic characters between syntopic *R. okaloosae* and *R. clamitans* (Moler, 1985), and the lack of monophyly of *R. okaloosae* haplotypes suggest recent speciation as a result of strong divergent selection on phenotypic characters, with incomplete lineage sorting at mtDNA (the average *cyt b* pairwise divergence between *R. clamitans* and *R. okaloosae* in the CPA is 0.007). Alternatively, *R. clamitans* and *R. okaloosae* may have diverged long ago, but under-

gone matrilineal introgression, perhaps following a pronounced reduction in population size in *R. okaloosae*. However, a definitive statement on the origin of *R. okaloosae* cannot be made without the examination of independent markers. To date, we have sequenced only a small number of *R. okaloosae* mtDNA sequences, and although all have shared *R. clamitans* haplotypes, we cannot yet rule out that rare divergent *R. okaloosae*-specific haplotypes do exist.

4.1. Patterns and implications of mtDNA evolution

Most protein coding regions are thought to be under the influence of negative selection because changes to well-adapted protein functions are less likely to persist unless they are adaptive (Hughes, 1999; Kreitman, 2000). Accordingly, non-adaptive, non-synonymous changes should be found predominantly at the tips of gene trees, where they are relatively young mutations that have yet to experience negative selection (Zink, 2005). The distribution of replacement substitutions (as indicated by non-significant neutral substitution rates, Fig. 2) in both *R. clamitans* and *R. catesbeiana* supports our interpretation of weak negative selection.

In both *R. clamitans* and *R. catesbeiana*, evidence of adaptive variation is detected at the intra- interspecific boundary of *cyt b*. In *R. catesbeiana*, non-directional selection is indicated deep in the diversification of the west lineage, while directional selection is apparent in the Louisiana *R. clamitans* *cyt b* haplotypes. This pattern underscores the concerns that mtDNA should not be assumed neutral and that care should be taken to reconcile the scale of study (e.g., population level, regional, or range-wide) with the signal and pattern of selection of the chosen marker at that particular scale. In our case, it appears that the mtDNA loci under investigation are evolving according to the nearly neutral model for much of the more recent phylogenetic history of the group, and therefore is not likely to have an important influence on our results.

Our M–K tests show gene-specific differences between ratios of variability in replacement and synonymous changes. For most comparisons the overall ratios of synonymous and nonsynonymous substitutions were not significant, suggesting that variation at the intraspecific level (and interspecifically at ND2) is due to genetic drift rather than natural selection. However, differences between *R. clamitans* and *R. catesbeiana* at *cyt b* were significant, although it is unclear from M–K test whether this is caused by processes (e.g., selection) occurring in one or both species. Because mtDNA lacks recombination and is inherited as a single locus variation representing the arrival and spread of advantageous mutations in *cyt b* should result in deviations from neutrality at ND2 as a result of genetic hitchhiking. However, if weakly deleterious mutations persist only for short periods, they are not likely to contribute to interspecific divergence (Ohta, 1992) across linked loci. Weakly deleterious non-synonymous mutations have been proposed as the determinants of a similar polymorphism pattern in *Drosophila* (Ohta, 1992), mice (Nachman et al., 1994) and primates (Wise et al., 1998; Nachman et al., 1996). Both *R. clamitans* and *R. catesbeiana* show fewer replacement substitutions than is predicted by neutral theory (25–30% of branches for ND2 in both *R. clamitans* and *R. catesbeiana*, and at *cyt b* for *R. catesbeiana*). In the case of *cyt b*, the proportion of significant values is considerably greater in *R. clamitans* (over 50%), suggesting that the paucity of replacement substitutions in *R. clamitans* is driving the significance in the M–K test, and possibly causing the relatively shallow lineage divergence compared to *R. catesbeiana*.

5. Conclusions

Evidence of varying patterns of non-neutral evolution in *cyt b* and ND2 common in both *R. clamitans* and *R. catesbeiana* highlights

the importance of understanding the molecular evolution of the mitochondrial genome, particularly with respect to its widespread use as a tool for demographic reconstruction (Bazin et al., 2006). Despite this seemingly dire observation, our point is not that mtDNA is an inappropriate marker, but that understanding temporal or demographic patterns require careful consideration of major assumptions, in this case neutrality of genetic markers. This is the case for any marker, and there is little evidence that the increasingly common nuclear genes used in inter- and some intra-specific phylogenies are any less susceptible to deviations from assumptions than is mtDNA. However, the relatively greater utility of mtDNA compared with nDNA in reconstructing phylogeographic history in some taxa where they have been compared seems clear (e.g., Zink and Barrowclough, 2008).

The geographic distributions of lineages and timing of differentiation within each species differ considerably. This may reflect differing responses to shared temporal phenomena (e.g., allopatric isolation in different areas during the Pliocene, and differences in Pleistocene post-glacial dispersal pattern). Insofar as we can accurately infer population and biogeographic history from a single molecular marker and accommodate selection in our interpretations, the patterns described here suggests that these two species have undergone dramatically different geographic patterns of population isolation and expansion, though in both cases from similarly small ancestral populations and at broadly similar time periods. If the assumption of approximate evolutionary rates in ecologically similar species (Thorne and Kishino, 2005) is false, then our estimates may be spurious.

Acknowledgments

We thank J.J. Apodaca, R. Birkhead, D. Bishop, J. Gilhen, S. Loughheed, W. Meshaka, B. Miller, P. Moler, G. Nelson, S. Richter, E. Wild, and J. Lee-Yaw, for *Rana* samples; F. Chen for assistance with sequencing; and S. Loughheed and M. Vallianatos for constructive criticism on earlier drafts of the manuscript. Support for J.D.A. came from a Natural Science and Engineering Research Council Postdoctoral Fellowship (Canada). Lab work was funded in part by NSF Grant DEB 0343526 (to K.Z.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ymp.2008.06.010.

References

- Al-Rabab'ah, M.A., Williams, C.G., 2002. Population dynamics of *Pinus taeda* L. based on nuclear microsatellites. *For. Ecol. Manage.* 163, 263–271.
- Arbogast, B.S., Edwards, S.V., Wakeley, J., Beerli, P., Slowinski, J.B., 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Ann. Rev. Ecol. Syst.* 33, 707–740.
- Austin, J.D., Loughheed, S.C., Moler, P., Boag, P.T., 2003. Phylogenetics, zoogeography, and the role of vicariance and dispersal in the evolution of the *Rana catesbeiana* (Anura: Ranidae) species group. *Biol. J. Linn. Soc.* 80, 601–624.
- Austin, J.D., Loughheed, S.C., Boag, P.T., 2004a. Discordant temporal and geographic patterns in maternal lineages of eastern North American frogs, *Rana catesbeiana* (Ranidae) and *Pseudacris crucifer* (Hylidae). *Mol. Phylogenet. Evol.* 32, 799–816.
- Austin, J.D., Loughheed, S.C., Boag, P.T., 2004b. Controlling for the effects of history and nonequilibrium conditions in gene flow estimates in northern bullfrog (*Rana catesbeiana*) populations. *Genetics* 168, 1491–1506.
- Avice, J.C., 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge.
- Avice, J.C., Walker, D., Johns, G.C., 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc. R. Soc. Lond. B* 265, 1707–1712.
- Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13, 729–744.
- Barton, N.H., Etheridge, A.M., 2004. The effect of selection on genealogies. *Genetics* 166, 1115–1131.
- Bazin, E., Glemin, S., Galtier, N., 2006. Population size does not influence mitochondrial genetic diversity in animals. *Science* 312, 570–572.

- Bernatchez, L., Wilson, C.C., 1998. Comparative phylogeography of nearctic and paleartic freshwater fishes. *Mol. Ecol.* 7, 431–452.
- Bridle, J.R., Pedro, P.M., Butlin, R.K., 2004. Habitat fragmentation and biodiversity: testing for the evolutionary effects of refugia. *Evolution* 58, 1394–1396.
- Brooks Jr., G.R., 1964. An analysis of the food habits of the bullfrog, *Rana catesbeiana*, by body size, sex, month and habitat. *Va. J. Sci.* 15, 173–186.
- Brown, W.M., 1985. The mitochondrial genome of animals. In: MacIntyre, R.J. (Ed.), *Molecular Evolutionary Genetics*. Plenum Press, New York, pp. 95–130.
- Brunsfeld, S.J., Sullivan, J., Soltis, D.E., Soltis, P.S., 2001. Comparative phylogeography of north-western North America: a synthesis. In: Silvertown, J., Antonovics, J. (Eds.), *Integrating Ecological and Evolutionary Processes in a Spatial Context*. Blackwell Science, Oxford, pp. 319–339.
- Burbrink, F.T., Lawson, R., Slowinski, J.B., 2000. Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept. *Evolution* 54, 2107–2118.
- Caccone, A., Milinkovitch, M.C., Sbordoni, V., Powell, J.R., 1997. Mitochondrial DNA rates and biogeography in European newts (genus *Euproctus*). *Syst. Biol.* 46, 126–144.
- Camin, J.H., Sokal, R.L., 1965. A method to deduce branching sequences in phylogeny. *Evolution* 19, 311–326.
- Carstens, B.C., Brunsfeld, S.J., Demboski, J.R., Good, J.M., Sullivan, J., 2005. Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: Hypothesis testing within a comparative phylogeographic framework. *Evolution* 59, 1639–1652.
- Clark, A.M., Moler, P.E., Possardt, E.E., Savitzky, A.H., Brown, W.S., Bowen, B.W., 2003. Phylogeography of the timber rattlesnake (*Crotalus horridus*) based on mtDNA sequences. *J. Herpetol.* 37, 145–154.
- Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9, 1657–1660.
- Conant, R., Collins, J.T., 1998. *A Field Guide to Reptiles & Amphibians: Eastern and Central North America*. Houghton Mifflin Company, New York.
- Creevey, C.J., McInerney, J.O., 2002. An algorithm for detecting directional and non-directional positive selection, neutrality and negative selection in protein coding DNA sequences. *Gene* 300, 43–51.
- Creevey, C.J., McInerney, J.O., 2003. Crann: detecting positive selection in protein coding DNA sequences. *Bioinformatics* 19, 1726.
- Dubois, A., 1992. Notes sur la classification des Ranidae (Amphibiens Anoures). *Bull. Soc. Linn. Lyon* 61, 305–352.
- Durham, L., Bennett, G.W., 1963. Age, growth, and homing in the bullfrog. *J. Wildl. Manage.* 27, 107–123.
- Gillespie, J.H., 2001. Is the population size of a species relevant to its evolution? *Evolution* 55, 2161–2169.
- Goebel, A.M., Donnelly, J.M., Atz, M.E., 1999. PCR primers and amplification methods for 12S ribosomal DNA, the control region, cytochrome oxidase I, and cytochrome *b* in bufonids and other frogs, and an overview of PCR primers which have amplified DNA in amphibians successfully. *Mol. Phylogenet. Evol.* 11, 163–199.
- Hamilton Jr., W.J., 1948. The food and feeding behavior of the green frog, *Rana clamitans* Latreille, in New York State. *Copeia* 1948, 203–207.
- Hecnar, S.J., McCloskey, R.T., 1997. Changes in the composition of a ranid frog community following bullfrog extinction. *Am. Midl. Nat.* 137, 145–150.
- Hewitt, G.M., 1999. Post-glacial recolonization of European biota. *Biol. J. Linn. Soc.* 68, 87–112.
- Hey, J., Nielsen, L., 2004. Multilocus methods for estimating population sizes, migration rates, and divergence times, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167, 747–760.
- Hey, J., Nielsen, L., 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. USA* 104, 2785–2790.
- Hillis, D.M., Davis, S.K., 1986. Evolution of ribosomal DNA: fifty million years of recorded history in the frog genus *Rana*. *Evolution* 40, 1275–1288.
- Hillis, D.M., Wilcoz, T.P., 2005. Phylogeny of the New World true frogs (*Rana*). *Mol. Phylogenet. Evol.* 34, 299–314.
- Ho, S.Y.W., Phillips, M.J., Cooper, A., Drummond, A.J., 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* 22, 1561–1568.
- Howes, B.J., Lindsay, B., Lougheed, S.C., 2006. Range-wide phylogeography of a temperate lizard, the five-lined skink (*Eumeces fasciatus*). *Mol. Phylogenet. Evol.* 40, 183–194.
- Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.
- Huelsenbeck, J.P., Hillis, D.M., Jones, R., 1996. Parametric bootstrapping in molecular phylogenetics: applications and performance. In: Ferraris, J.D., Palumbi, S.R. (Eds.), *Molecular Zoology: Advances, Strategies, and Protocols*. Wiley-Liss, New York, pp. 19–45.
- Hughes, A.L., 1999. *Adaptive Evolution of Genes and Genomes*. Oxford University Press, Oxford.
- Johnson, N.K., Cicero, C., 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution* 58, 1122–1130.
- Klicka, J., Zink, R.M., 1999. Pleistocene effects on North American songbird evolution. *Proc. R. Soc. Lond. B* 266, 695–700.
- Kreitman, M., 2000. Methods to detect selection in populations with applications to the human. *Annu. Rev. Genomics Hum. Genet.* 1, 539–559.
- Lapointe, F.J., Rissler, L.J., 2005. Congruence, consensus, and the comparative phylogeography of codistributed species in California. *Am. Nat.* 166, 290–299.
- Leache, A.D., Reeder, T.W., 2002. Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): a comparison of parsimony, likelihood, and Bayesian approaches. *Syst. Biol.* 51, 44–68.
- Li, W.H., 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.* 36, 96–99.
- Macey, J.R., Strasburg, J.L., Brisson, J.A., Vredenburg, V.T., Jennings, M., Larson, A., 1998. Molecular phylogenetics of western North American frogs of the *Rana boylei* species group. *Mol. Phylogenet. Evol.* 19, 131–143.
- Martin, A.P., Palumbi, S., 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl. Acad. Sci. USA* 90, 4087–4091.
- Martof, B., 1956. Growth and development of the green frog, *Rana clamitans*, under natural conditions. *Am. Midl. Nat.* 55, 101–117.
- McDonald, J.H., Kreitman, M., 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351, 652–654.
- Mecham, J.S., 1954. Geographic variation in the green frog, *Rana clamitans* Latreille. *Tex. J. Sci.* 6, 1–24.
- Moler, P., 1985. A new species of frog (Ranidae: *Rana*) from northwestern Florida. *Copeia* 1985, 379–383.
- Moriarty, E.C., Cannatella, D.C., 2004. Phylogenetic relationships of the North American chorus frogs (Pseudacris: Hylidae). *Mol. Phylogenet. Evol.* 30, 409–420.
- Moritz, C., Faith, D.P., 1998. Comparative phylogeography and the identification of genetically divergent areas for conservation. *Mol. Ecol.* 7, 419–429.
- Moritz, C., Schneider, C.J., Wake, D.B., 1992. Evolutionary relationships within *Ensatina eschscholtzii* complex confirm the ring species interpretation. *Syst. Biol.* 41, 273–291.
- Mueller, R.L., 2006. Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Syst. Biol.* 55, 289–300.
- Nachman, M.W., Boyer, S.N., Aquadro, C.F., 1994. Non-neutral evolution at the mitochondrial ND3 gene in mice. *Proc. Natl. Acad. Sci. USA* 91, 6364–6368.
- Nachman, M.W., Brown, W.M., Stoneking, M., Aquadro, C.F., 1996. Nonneutral mitochondrial DNA variation in humans and chimpanzees. *Genetics* 142, 953–963.
- Neuhauser, C., Krone, S.M., 1997. The genealogy of samples in models with selection. *Genetics* 145, 519–534.
- Nielsen, R., Wakeley, J., 2001. Distinguishing migration from isolation. A Markov chain Monte Carlo approach. *Genetics* 158, 885–896.
- Ohta, T., 1992. The nearly neutral theory of molecular evolution. *Ann. Rev. Ecol. Syst.* 23, 263–286.
- Plötner, J., Ohst, T., Bohme, W., Schreiber, R., 2001. Divergence in mitochondrial DNA of near eastern water frogs with special reference to the systematic status of Cypriot and Anatolian populations. *Amph. Rept.* 22, 397–412.
- Posada, D., Crandall, K.A., 1998. ModelTest: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Posada, D., Crandall, K.A., 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol. Evol.* 16, 37–45.
- Rambaut, A., Grassly, N.C., 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput. Appl. Biosci.* 13, 235–238.
- Rand, D.M., Kahn, L.M., 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice and humans. *Mol. Biol. Evol.* 13, 735–748.
- Rowe, K.C., Heske, E.J., Paige, K.N., 2006. Comparative phylogeography of eastern chipmunks and white-footed mice in relation to the individualistic nature of species. *Mol. Ecol.* 15, 4003–4020.
- Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analysis by the coalescent and other methods. *Bioinformatics* 19, 2496–2497.
- Ryan, M.J., 1980. The reproductive behavior of the bullfrog (*Rana catesbeiana*). *Copeia* 1980, 108–114.
- Schroeder, E.E., 1976. Dispersal and movement of newly transformed green frogs, *Rana clamitans*. *Am. Midl. Nat.* 95, 471–474.
- Schroeder, E.E., Baskett, T.S., 1968. Age estimation, growth rates, and population structure in Missouri bullfrogs. *Copeia* 1968, 583–592.
- Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., Soltis, P.S., 2006. Comparative phylogeography of unglaciated eastern North America. *Mol. Ecol.* 15, 4261–4293.
- Steele, C.A., Storfer, A., 2007. Phylogeographic incongruence of codistributed amphibian species based on small differences in geographic distribution. *Mol. Phylogenet. Evol.* 43, 468–479.
- Swofford, D.L., 2002. PAUP4*: Phylogenetic Analysis Using Parsimony (*and other methods), ver. 4. Sinauer Associates, Sunderland, MA.
- Taberlet, P., Fumagalli, L., Wust-Saucy, A.G., Cosson, J.F., 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* 7, 453–464.
- Tajima, F., 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics* 105, 437–460.
- Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.
- Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA 4: molecular evolutionary genetic analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596–1599.
- Tan, A.-M., Wake, D.B., 1995. MtDNA phylogeography of the California newt, *Taricha torosa* (Caudata, Salamandridae). *Mol. Phylogenet. Evol.* 4, 383–394.

- Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonucleases mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 13, 597–601.
- Thorne, J.L., Kishino, H., 2005. Estimation of divergence times from molecular sequence data. In: Nielsen, R. (Ed.), *Statistical Methods in Molecular Evolution*. Springer, New York, pp. 233–256.
- Wells, K.D., 1977. Territoriality and male mating success in the green frog (*Rana clamitans*). *Ecology* 58, 750–762.
- Werner, E.E., McPeck, M.A., 1994. Direct and indirect effects of predators on two anuran species along an environmental gradient. *Ecology* 75, 1368–1382.
- Willis, Y.L., Moyle, D.L., Baskett, T.S., 1956. Emergence, breeding, hibernation, movements, and transformations of the bullfrog, *Rana catesbeiana* in Missouri. *Copeia* 1956, 596–635.
- Wise, C.A., Sraml, M., Easteal, S., 1998. Departure from neutrality at the mitochondrial NADH dehydrogenase subunit 2 gene in humans, but not in chimpanzees. *Genetics* 148, 409–421.
- Yang, Z., Bielawski, J.P., 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* 15, 496–503.
- Zamudio, K.R., Savage, W.K., 2003. Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution* 57, 1631–1652.
- Zink, R.M., 2005. Natural selection on mitochondrial DNA in *Parus* and its relevance for phylogeographic studies. *Proc. R. Soc. Lond. B* 272, 71–78.
- Zink, R.M., Barrowclough, G.F., 2008. Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* 17, 2107–2121.
- Zink, R., Drovetski, M.S.V., Rohwer, S., 2006. Selective neutrality of mitochondrial ND2 sequences, phylogeography and species limits in *Sitta europaea*. *Mol. Phylogenet. Evol.* 40, 679–686.