Ichthyology & Herpetology 110, No. 3, 2022, 602–617

Genomic Data Clarify *Aquarana* Systematics and Reveal Isolation-by-Distance Dominates Phylogeography of the Wide-Ranging Frog *Rana clamitans*

Daniel J. MacGuigan[1,¶], Genevieve G. Mount[2,3,¶], Gregory J. Watkins-Colwell[4], Thomas J. Near[4,5], and Max R. Lambert[6]

Wide-ranging species often span geographic dispersal barriers, providing opportunity for divergence via genetic drift or natural selection. Such conditions can be favorable for speciation, and wide-ranging taxa are frequently subdivided into multiple species by modern molecular studies. However, with wide-ranging species, it is important to explicitly test for isolation-by-distance (IBD), which can produce continuous genetic variation that may be misinterpreted as discrete population structure or even distinct species. Here we examine the Green Frog (*Rana clamitans*), a wide-ranging species of the *Aquarana* clade distributed across much of North America. The broader phylogenetic context for *R. clamitans* within *Aquarana* is poorly understood, particularly its relationship with the closely related and IUCN Vulnerable Florida Bog Frog (*Rana okaloosae*). Additionally, although phenotypic variation within *Rana clamitans* previously merited description of two subspecies, patterns of genetic diversity remain unclear. Using genome-wide ddRAD markers, we largely resolve relationships within *Aquarana* and unambiguously identify *R. okaloosae* as the sister lineage to *R. clamitans*. Despite mtDNA introgression, we find little genomic evidence of hybridization between *R. okaloosae* and *R. clamitans*. Within *R. clamitans* there are two well-supported and geographically divided clades, which are identified as distinct species by a multispecies coalescent-based approach. However, the two clades exhibit low genealogical divergence. Population genetic analyses reveal that genetic variation within *R. clamitans* is best described by a pattern of IBD rather than independently evolving lineages. We conclude that *R. clamitans* is indeed a single species and the subspecies concept is of limited use. Our analysis demonstrates the importance of understanding continuous genetic variation when delimiting lineages and highlights the power of combining population genetic and phylogenetic perspectives to describe diversity within wide-ranging taxa.

IDENTIFYING boundaries and relationships among populations and species is essential to how we organize biodiversity and understand evolution and ecology. Inherently, the purpose of delimiting species is to designate evolutionarily significant units, meaningful groups of individuals and populations that vary in terms of their role within a biological community and which can be placed in a larger phylogenetic context (de Queiroz, 2007; Patten and Remsen, 2017; Hillis, 2019). Historically, intraspecific diversity—phenotypic and/or genotypic—may have been classified as subspecies. However, for many clades, subspecies are becoming rare as taxonomists either elevate subspecies to full species or eliminate subspecies designations (e.g., Kassler et al., 2002; Near et al., 2011; Torstrom et al., 2014; MacGuigan et al., 2017; Chambers and Hillis, 2020; Chan et al., 2020; Hillis, 2020; Marshall et al., 2021). Delimiting variation within a species or between close relatives provides the crucial foundation for characterizing broader phylogenetic relationships.

Lineage boundaries within widespread species are particularly difficult to delimit, as continuous genetic variation may be erroneously identified as discrete species (e.g., Hillis, 2020; Burbrink and Ruane, 2021; Hillis et al., 2021). Wide-ranging taxa often contain morphological and genetic variation due to isolation-by-distance (IBD) or variation in habitat and ecology (Niemiller et al., 2011; Dincă et al., 2019; Chambers and Hillis, 2020). Historical approaches to studying wide-ranging taxa, including morphological and early molecular datasets with few genetic loci, often had limited power to infer evolutionary history at recent time scales. While mitochondrial DNA (mtDNA) has a long history of use in phylogeography, species delimitation, and shallow scale phylogenetics, it is prone to "capture" by closely related lineages (Chan and Levin, 2005; Good et al., 2008; Hillis, 2020). As large genomic datasets have become increasingly accessible for diverse groups of organisms, we can now use genome-wide information to uncover evolutionary histories for many clades (Dincă et al., 2019).

Here we scrutinize the semiaquatic Green Frog (*Rana clamitans*) through a genomic lens as a case study to illustrate challenges associated with describing intraspecific versus interspecific genetic variation in wide-ranging species. Occurring from eastern Texas north into Ontario, eastward through Nova Scotia, and south through Florida, *R. clamitans* is a widespread and abundant North American amphibian (AmphibiaWeb, 2020). *Rana clamitans* comprises two previ-

[1] Department of Biological Sciences, University at Buffalo, Buffalo, New York 14260; Email: dmacguig@buffalo.edu. Send correspondence to this address.
[2] Department of Biology, Utah State University, Logan, Utah 84322; Email: ggmount@gmail.com.
[3] Museum of Vertebrate Zoology, UC Berkeley, Berkeley, California 94720.
[4] Division of Vertebrate Zoology, Peabody Museum of Natural History, Yale University, New Haven, Connecticut 06511; Email: (GJWC) gregory.watkins-colwell@yale.edu.
[5] Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06511; Email: thomas.near@yale.edu.
[6] Science Division, Habitat Program, Washington Department of Fish & Wildlife, Olympia, Washington 98501; Email: lambert.mrm@gmail.com.
[¶] Co-first authors.
Submitted: 15 November 2021. Accepted: 16 June 2022. Associate Editor: B. L. Stuart.
 DOI: 10.1643/h2021129 Published online: 31 October 2022

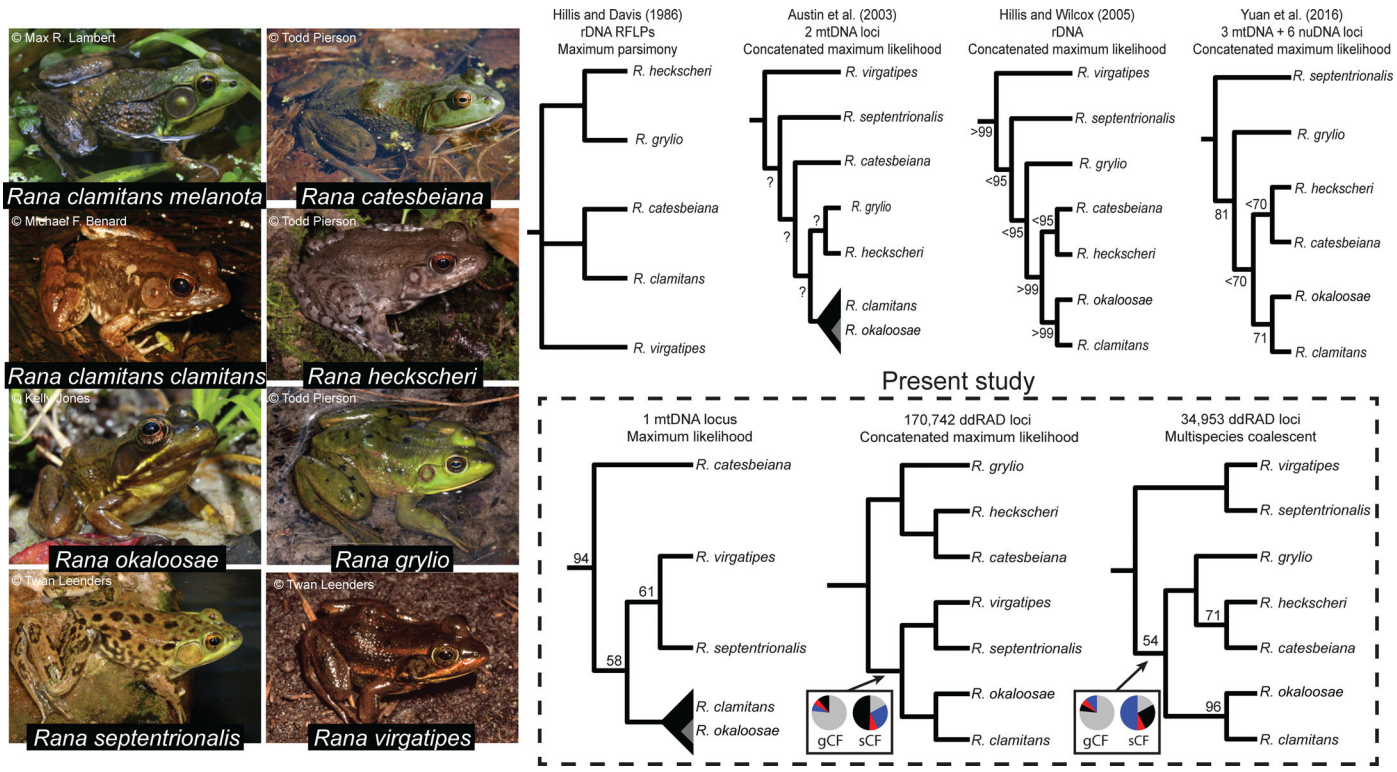

Fig. 1. Previously published phylogenetic relationships for the *Aquarana* clade and results from our analyses using a single mitochondrial locus and two approaches with ddRAD data. Nodes have 100% bootstrap support unless otherwise labeled. Support values were not reported in Hillis and Davis (1986) or Austin et al. (2003). Pie charts indicate site concordance factor (sCF) or gene concordance factor (gCF) support for the only topological difference between the ddRAD species tree and the ddRAD concatenated tree. For each pie chart, black indicates agreement with the displayed tree, red and blue indicate disagreement, and gray indicates no support or noise.

ously identified subspecies: the Bronze Frog (*R. c. clamitans*) and the Green Frog (*R. c. melanota*; Fig. 1; Crother, 2017; Robert Powell, pers. comm.). Despite its abundance and broad distribution, the taxonomy of *R. clamitans* has remained unclear for over a century. Based largely on qualitative evidence, Rhoads (1895) described two subspecies of *R. clamitans*, suggesting that *R. c. clamitans* is a smaller subspecies with a more uniform green coloration that is distributed in the southern portion of the range and *R. c. melanota* as a larger subspecies with darker pigmentation and a northern distribution. However, these names went unused until Mecham (1954) performed a qualitative assessment of coloration and body size to resurrect the two subspecies (Fig. 1). The geographic boundary of Mecham's (1954) subspecies delimitation roughly followed the edge of the southern coastal plain (Fig. 2) and relied on coloration and relative disparity in body size, though there were no consistently diagnostic phenotypic differences between the two subspecies.

Despite the questionable nature of *Rana c. clamitans* and *R. c. melanota*, recognition of these two subspecies was commonly accepted by biologists until the late 2010s (AmphibiaWeb, 2020). Early molecular analyses using mtDNA gene sequences did not resolve the subspecies as reciprocally monophyletic groups, but rather suggested modest support for three lineages within *R. clamitans* (Austin et al., 2003; Austin and Zamudio, 2008). None of these clades corresponded to Mecham's (1954) proposed subspecies. These molecular phylogenies cast doubt on Mecham's (1954) long-accepted subspecies delimitations in *R. clamitans*, leading to the exclusion of these subspecies in the most recent standard naming guides for North American herpetofauna (Crother, 2017; Stebbins and McGinnis, 2018). However, phylogeographic patterns within *R. clamitans* using more comprehensive genome-wide markers remain unexplored.

Of additional interest is the evolutionary relationship between *R. clamitans* and a close relative, the Florida Bog Frog (*R. okaloosae*), an IUCN-listed Vulnerable species (Fig. 1). In contrast to the large range of *R. clamitans*, *R. okaloosae* is restricted to a minuscule range in just two Florida counties, with most known populations occurring on Eglin Air Force Base (Moler, 1985; Austin et al., 2003; Austin and Zamudio, 2008; Gorman et al., 2009; Priestley et al., 2010; Gorman and Haas, 2011). Larvae and adults of *R. okaloosae* and *R. clamitans* are phenotypically and ecologically distinguishable (Moler, 1985; Gorman et al., 2009; Priestley et al., 2010; Gorman and Haas, 2011). However, previous molecular analyses suggest these two species are not distinct, as mitochondrial DNA phylogenies indicate *R. okaloosae* is nested within *R. clamitans*, perhaps due to hybridization (Austin et al., 2003; Austin and Zamudio, 2008). Despite this, *R. okaloosae* has consistently been considered a distinct species from *R. clamitans*, likely due to the obvious morphological and ecological differences. Analyses of genome-wide nuclear markers offer a powerful approach to investigate the phylogenetic relationships, possible ongoing hybridization, and species status of *R. okaloosae* and syntopic *R. clamitans* (Austin and Zamudio, 2008).

Beyond potential hybridization with *R. okaloosae*, the systematics of *R. clamitans* and other closely related species also remain notably unresolved. Within the true frogs (Ranidae), *Rana* is the most species-rich genus and is well represented in the Americas (Hillis and Wilcox, 2005; Yuan et al., 2016; AmphibiaWeb, 2020). Though prior work proposed splitting North American *Rana* into two genera—*Rana* and *Lithobates* (Frost et al., 2006, 2008; Wiens, 2007)—doing so promoted taxonomic instability and resulted in paraphyly (Pauly et al., 2009; Yuan et al., 2016) and so we refrain from using *Lithobates*. The *Rana* subgenus *Aquarana* includes seven recognized species restricted to central and eastern North American (Fig. 1). *Aquarana* is strongly supported as monophyletic in molecular phylogenetic analyses (Fig. 1). Yet despite the application of a variety of molecular markers including restriction fragment length polymorphisms (Hillis and Davis, 1986), mitochondrial DNA (Austin et al., 2003; Hillis and Wilcox, 2005), and combined mitochondrial and nuclear DNA (Yuan et al., 2016), relationships within *Aquarana* remain unresolved. Genome-wide markers have promise to clarify questions at multiple scales of biodiversity, from intraspecific variation within *R. clamitans* to interspecific phylogenetic relationships in the *Aquarana* clade.

Here we leverage genomic data to investigate genetic diversity within *Rana clamitans*, determine if there is historical or ongoing hybridization between *R. clamitans* and the Vulnerable *R. okaloosae*, and infer phylogenetic relationships of the *Aquarana* clade. ddRAD sequencing provides tens to hundreds of thousands of variable loci throughout the genome and is increasingly common in phylogenetic and species delimitation studies, particularly for amphibians with large genomes (Dufresnes et al., 2018; McCartney-Melstad et al., 2018; Reyes-Velasco et al., 2018; Weisrock et al., 2018). Our findings here reveal the phylogeographic history of *R. clamitans* and clarify relationships among species in the *Aquarana* clade. Additionally, we provide the framework for future investigations of potential introgression between *R. clamitans* and *R. okaloosae*. Finally, this study highlights some challenges and pitfalls of differentiating between intraspecific and interspecific genetic variation within widely distributed species.

MATERIALS AND METHODS

Sample collection and sequencing.—We acquired 90 tissue samples and DNA extractions covering all seven species of *Aquarana*: *Rana clamitans*, *R. okaloosae*, *R. septentrionalis*, *R. virgatipes*, *R. catesbeiana*, *R. heckscheri*, *R. grylio* (Table S1; see Data Accessibility). Note that the wood frog (*R. sylvatica*) is occasionally included in *Aquarana* (Frost, 2022), but its position within *Rana* has been difficult to resolve (AmphibiaWeb, 2020). However, prior phylogenetic analyses consistently resolve a clade containing all species of *Aquarana* exclusive of *R. sylvatica* (Hillis and Davis, 1986; Austin et al., 2003; Hillis and Wilcox, 2005; Yuan et al., 2016). Thus, we utilized *R. sylvatica* and another distant relative, *R. sphenocephala*, as phylogenetic outgroups. Sampling focused on *R. clamitans*, with 75 individuals from 39 range-wide localities.

Following Yale IACUC protocol 2015-110681, we collected and euthanized frogs in MS-222 prior to dissecting muscle or liver tissue, which were preserved in 95% ethanol. Tissue samples were also acquired from the following institutions: the Illinois Natural History Survey, the New Brunswick Museum Tissue Collection, and the Field Museum of Natural History. Additional samples were generously provided by Dr. James Austin at the University of Florida. Genomic DNA was extracted from tissues using a DNeasy Qiagen Blood and Tissue Kit. We cleaned DNA extractions using ethanol precipitation. 3M sodium acetate (pH = 5.2) was added equal to 10% of the total volume of the DNA extraction, followed with 100% ethanol equal to 2.5 times the total volume of DNA. Extractions were incubated for ten minutes at –80°C, then centrifuged for 30 minutes at 8,000 RCF. The supernatant was poured off and the DNA pellet was washed with 250 μL of cold 70% ethanol. Samples were centrifuged again for five minutes at 8,000 RCF and the supernatant was poured off. We allowed the pellet to air dry for ~15 minutes, then resuspended with the desired amount of DNase-free water.

We performed a modified version of the double digest restriction site-associated DNA (ddRAD) sequencing protocol outlined in Peterson et al. (2012) and Poland et al. (2012). We digested 200 ng of DNA from each sample for eight hours at 37°C using the restriction enzymes PstI and SbfI. Samples were visually checked using a 2.5% agarose gel to ensure complete digestion. We then ligated custom adapters using T4 DNA ligase and an incubation period of three hours at 22°C. The adapters contained a set of 96 unique 8–10 bp barcodes. Barcode sequences were generated with at least two mutational differences between any pair of barcodes to avoid misassignment during demultiplexing. Samples were then pooled and cleaned using a QIAquick PCR purification kit. We performed 12 rounds of polymerase chain reaction (PCR), each cycle with 30 seconds at 98°C, 30 seconds at 62°C, and 30 seconds at 72°C. Following the 12 cycles, we held for ten minutes at 17°C. After PCR, sets of 96 samples with unique barcodes were pooled to perform a 300–500 base pair size selection using a Blue Pippen 2% agarose gel cassette. Size selection was confirmed using an ABI Bioanalyzer High Sensitivity DNA assay. Size selection and fragment analysis was performed at the Yale DNA Analysis Facility on Science Hill. We sequenced the 96-sample genomic library using one lane of Illumina HiSeq 4000 with 100 base pair single-end reads at the University of Oregon Genomics and Cell Characterization Core Facility. We inspected raw read quality using FastQC (Andrews, 2010).

We used the published genome of *Rana catesbeiana* (GenBank GCA_002284835.2) to assemble the ddRAD dataset using iPyrad v.0.7.30 (Eaton and Overcast, 2020). After demultiplexing the Illumina reads with iPyrad step 1, we trimmed raw reads using CutAdapt (Martin, 2011). Reads were trimmed to 100 bp and three rounds of trimming were used to remove adapter contamination and restriction cut sites ("-b TGCAG", "-a TGCAG", and "-a CCGAGATCGGAA GAGC"). Trimmed reads were then used as input for iPyrad steps 2–7 (see Table S2 for assembly parameters; see Data Accessibility).

In addition to ddRAD sequencing, we also amplified the mitochondrial cytochrome b (*cytb*) gene using PCR with previously described primers MVZ15-L (Moritz et al., 1992) and MVZar-H (Goebel et al., 1999) and reaction conditions (Austin and Zamudio, 2008). Amplified double-stranded DNA was purified using 20% polyethylene glycol DNA precipitation. Both the forward and reverse DNA strands were sequenced by the Yale Keck DNA Sequencing Facility using Big Dye Terminator Reaction Kits (Applied Biosystems, Foster City, CA). Sequencing was carried out on an Applied

Biosystems 3730xL (Applied Biosystems, Foster City, CA). We trimmed and assembled reads into contiguous sequences using Geneious v.8.0.5 (https://www.geneious.com) and aligned sequences using the program MUSCLE v.3.8.425 (Edgar, 2004). We checked each alignment of protein-coding sequences by eye for erroneous insertions or deletions.

Phylogenetic analyses.—We performed separate phylogenetic analyses of the *cytb* and the concatenated ddRAD datasets using IQ-Tree v.1.6.12 (Nguyen et al., 2015). To determine the impact of missing data on phylogenies inferred using ddRAD data, we used several datasets with different minimum samples-per-locus thresholds: 4/90 samples (170,742 loci, 86% missing data), 20/90 samples (34,953 loci, 63% missing data), and 40/90 (9,584 loci, 47% missing data). We used a GTR + gamma nucleotide substitution model for all ddRAD analyses. For analyses of the mtDNA data, we partitioned the alignment by codon position and performed substitution model optimization ("-m MFP+MERGE") with ModelFinder (Kalyaanamoorthy et al., 2017). We ran each tree search until 100 unsuccessful tree search iterations were completed. To assess topological support, we also performed 1,000 ultrafast bootstrap replicates (Hoang et al., 2018). In addition to concatenated phylogenetic analyses, we performed ddRAD species tree inference using the quartet-based method Tetrad v.0.9.13 (https://github.com/eaton-lab/tetrad). We used the most complete ddRAD dataset (40/90 individuals), analyzed all possible quartets, and performed 100 bootstrap replicates.

To examine phylogenetic discordance in the dataset, we estimated gene and site concordance factors using IQ-Tree. Gene concordance factors represent the proportion of gene trees that contain a given branch, while site concordance factors represent the proportion of SNPs supporting a branch. First, maximum likelihood gene trees were inferred for 31,665 ddRAD loci that contained at least 20 individuals and at least one parsimony informative site. IQ-Tree was then used to summarize the gene tree distribution, sampling 500 random quartets around each internal branch to estimate site concordance factors. We report site and gene concordance factors for all interspecific branches in the maximum likelihood ddRAD phylogeny.

We used the program TreeMix v.1.13 as implemented in iPyrad v.0.9.14 (Pickrell and Pritchard, 2012; Eaton and Overcast, 2020) to infer the history of gene flow among species of *Aquarana* with specific focus on *Rana okaloosae* and *R. clamitans*. We analyzed three different datasets. The first dataset contained *R. catesbeiana*, *R. septentrionalis*, *R. okaloosae*, the northeastern *R. clamitans* clade, and the southwestern *R. clamitans* clade (Table S3; see Data Accessibility). Since TreeMix relies on allele frequency data, we excluded *R. virgatipes*, *R. heckscheri*, and *R. grylio*, as these species were represented by only a single individual. The second and third datasets contained only *R. okaloosae* and *R. clamitans*. *Rana clamitans* was divided into four lineages using either the maximum likelihood (ML) phylogenetic tree or the LEA genetic clustering results to guide individual assignment (Table S3; see Data Accessibility). For each dataset, we used VCFTools v.0.1.15 (Danecek et al., 2011) to retain only biallelic sites with minor allele counts >1. Additionally, sites were required to be present in at least 10% of individuals for each lineage. For the TreeMix analyses, we allowed between one and five migration edges. Since different sets of SNPs may support different network topologies, we ran TreeMix using 100 data subsets containing one random SNP per 10 kbp genomic window. We summarized the results using a custom R script to extract the taxa and migration weights for each inferred migration edge supported by >10% of the SNP subsets.

Population genetic analyses.—We assembled several sets of SNPs using different filtering thresholds for missing data (Table S4; see Data Accessibility). Starting with the VCF file created by iPyrad, VCFTools was used to retain only biallelic sites (–max-alleles 2) and sites with less than 50%, 40%, or 30% missing data (–max-missing 0.5, 0.6, or 0.7). IBD analyses were performed using each of these datasets, which contained between 64,149 and 9,459 SNPs (Table S4; see Data Accessibility). To meet the assumption of independence among sites for genetic clustering analyses, we also thinned the dataset to include only a single random SNP per 10 kbp genomic window (–thin 10000). LEA and tess3R analyses were performed using each of the thinned datasets, which ranged from 8,498 to 1,224 SNPs (Table S4; see Data Accessibility).

Using these datasets, we tested for IBD using a Mantel test and linear models of genetic distance vs. geographic distance. We performed AICc model comparisons and model averaging between linear models containing interactions between geographic distance and clade comparisons, additive relationships between these covariates, univariate models, and a null model. First, we calculated geographic distances between sites using the *distm* function in the R package geosphere v.1.5 (Hijmans et al., 2016). We then computed Nei's genetic distance using the *dist.genpop* function in the R package adegenet v.2.1.1 (Jombart, 2008). We performed a Mantel test for IBD, comparing the observed correlation between geographic distance and genetic distance to the correlation from 10,000 permutations using the *mantel.randtest* function in the R package Ade4 v.1.7 (Dray and Dufour, 2007). Finally, we also performed a linear model of genetic distance vs. geographic distance using the *lm* function in R stats package v.3.6.0. We used the *dredge* function in the R package MuMIn (1.43.17; Bartón, 2009) to perform AICc model comparisons and model averaging between linear models containing interactions between geographic distance and clade comparisons, additive relationships between these covariates, univariate models, and a null model. Our phylogenetic and clustering analyses (see below) indicated two well-supported clades in *R. clamitans*. To reflect this in our IBD analysis, we coded all pairwise comparisons among *R. clamitans* as one of three levels: between the two clades (northeastern and southwestern) or within either the northeastern or southwestern clades. Adding this covariate allowed us to better understand the relative contribution of IBD and population structure in shaping intraspecific diversity of *R. clamitans*. A Tukey's *post hoc* test on the additive model (and assuming no slope differences among the three clade comparisons) was performed using the function *glht* in the R package *multcomp* v.1.4-15 (Hothorn et al., 2008).

Additionally, we used two genetic clustering methods, LEA v.3.0.0 (Frichot and François, 2015) and tess3R v.1.1.0 (Caye and Francois, 2016), to assess population structure by estimating ancestry coefficients. We performed genetic clustering analyses on a dataset of *R. clamitans* alone and a dataset including *R. clamitans* and *R. okaloosae*. Prior to

genetic clustering analyses, missing genotypes were imputed using the *impute* function (method = "random") from the R package LEA. First, we used the R package LEA to analyze both datasets with K = 1 through K = 20, ten replicates for each K value, and a maximum of 1,000 algorithm iterations per replicate. To examine the impact of the sNMF regularization, we ran three sets of analyses with alpha = 10, 100, or 500. We identified the optimal K value using cross-entropy scores. As LEA does not account explicitly account for geographic distance, we also ran tess3R v.1.1.0 (Caye and Francois, 2016), which incorporates geography into the clustering algorithm. Again, we ran both datasets with K = 1 through K = 20, ten replicates for each K value, and a maximum of 1,000 algorithm iterations per replicate. Tess3R cross-validation scores were used to determine the optimal K value.

Finally, to investigate the impact of missing data on genetic clustering, we performed two principal component analyses (PCA). We utilized a subset of 13 individuals as described in the Species Delimitation Analyses section below to assemble 7,872 ddRAD SNPs with no missing data. Additionally, we assembled a dataset of 77,945 SNPPs for the same 13 individuals but allowing <30% missing data per SNP. Datasets were assembled using iPyrad and VCFTools. Missing genotypes were imputed using the *tab* function (NA.method = "mean") from the R package adegenet. We then performed PCA using the *dudi.pca* function (scale = FALSE, scannf = FALSE, nf = 5) in the R package Ade4.

Species delimitation analyses.—To test whether population genetic structure in *Rana clamitans* is significant enough to constitute multiple species, we performed two sets of species delimitation analyses. Specifically, we tested whether the northeastern and southwestern clades of *R. clamitans* merit recognition as distinct species, rather than subspecies or clinal variation. We assembled a subset of the ddRAD data by strategically selecting five individuals from each clade that represented the breadth of genetic diversity and maximized the number of loci (Fig. 2). We also included all three samples of *R. okaloosae* in this dataset. Using iPyrad, we assembled a complete alignment requiring loci to be shared among all 13 individuals.

Both species delimitation analyses used BPP v.4.1.3 (Flouri et al., 2018) with a guide tree containing three tips: *R. okaloosae*, *R. clamitans* northeast, and *R. clamitans* southwest. To specify a biologically realistic prior on theta, we estimated nucleotide diversity using the *diversity.stats* function in the R package PopGenome v.2.7.5 (Pfeifer et al., 2014). Based on the average estimated genetic diversity across all three tips in the guide tree (0.003 mean substitutions per site), we specified an inverse gamma distribution with alpha = 3 and beta = 0.006. Given an estimated ranid frog mutation rate of 0.776E–9 (Sun et al., 2015) and a generation time of ~3 years (Shirose and Brooks, 1995), this is a broad prior with a mean N_e ~= 322,165 (sd = 1,288,660). Likewise, to specify a biological realistic prior on tau, we calculated the raw sequence distance between two individuals that capture the crown of *R. clamitans* and *R. okaloosae* using the *dist.dna* function in the R package ape v.5.4 (Paradis and Schliep, 2019). Based on a mean of 0.021 substitutions per site, we specified an inverse gamma prior with alpha = 5 and beta = 0.084, which translates to a wide distribution with a mean root age of 27 million years (sd = 16 million years). Together, both BPP priors are broad and relatively uninformative.

For the first species delimitation analysis, we used the BPP "10" algorithm, which performs reversible-jump Markov chain Monte Carlo (rjMCMC) to explore alternate delimitation models following a fixed guide tree. Specifically, we used BPP rjMCMC algorithm 1 (equal probabilities for rooted trees with alpha = 2 and m = 1), with four replicate analyses, 50,000 generations of burnin, and 1,000,000 MCMC generations, sampling every ten generations. We compared convergence among the replicate runs and summarized the posterior distribution for the number of supported species.

To complement the rjMCMC species delimitation analysis, we calculated the genealogical divergence index (GDI; Jackson et al., 2017) for each putative species. GDI values range between 0 and 1 and represent the degree of genetic divergence for a lineage. When there is no gene flow between two lineages, GDI is the probability that two intraspecific sequences coalesce before the species split when tracing the genealogy backward in time (Leaché et al., 2019). Following the suggestion of Jackson et al. (2017), we considered GDI < 0.2 as evidence of a single species, GDI > 0.7 as evidence of distinct species, and GDI values between 0.2 and 0.7 indicative of ambiguous delimitation. BPP assumes no gene flow between the lineages, which may result in an overestimate of GDI if gene flow is present (Jackson et al., 2017). Longer branch lengths (tau) and smaller effective population sizes (theta) produce larger GDI values. Using BPP, we estimated tau and theta under two guide trees. In the first guide tree, *R. clamitans* was divided into northeast and southwest species. In the second guide tree, *R. clamitans* was represented as a single species. For both sets of analyses, we used priors and MCMC conditions as specified above. We assessed mixing of MCMC runs by calculating ESS values and convergence by comparing across replicate runs. We used a custom R script to combine tau and theta estimates from replicate runs and calculate GDI.

RESULTS

ddRAD assembly.—All individuals in the ddRAD dataset (Table S1; see Data Accessibility) are represented by at least 778,000 reads, though there is significant variance in the number of reads per individual (mean = 3.7e6, sd = 2.1e6). We assembled alignments of 465,511 orthologous ddRAD loci mapping to the reference genome of *R. catesbeiana*. Using the initial alignments, we created filtered subassemblies for all subsequent analyses. The amount of genetic data in each assembly decreases rapidly as the minimum-samples-per-locus threshold increases (Table S5; see Data Accessibility). No ddRAD loci are shared by all samples, likely due to the genetic distance of the species of *Rana* in our datasets. In addition, we sequenced 911 base pairs of the partial coding region of the mtDNA gene *cytb*.

***Phylogeography of* Rana clamitans.**—We inferred ML ddRAD phylogenies using three datasets with different proportions of missing data. All phylogenies support two main clades within *Rana clamitans* (Figs. 2, S1; see Data Accessibility): a southwestern (SW) clade that is composed of specimens from Texas, Louisiana, Mississippi, and Arkansas (Figs. 2, S1; see Data Accessibility) and a northeastern (NE) clade that contains all other specimens of *R. clamitans* including

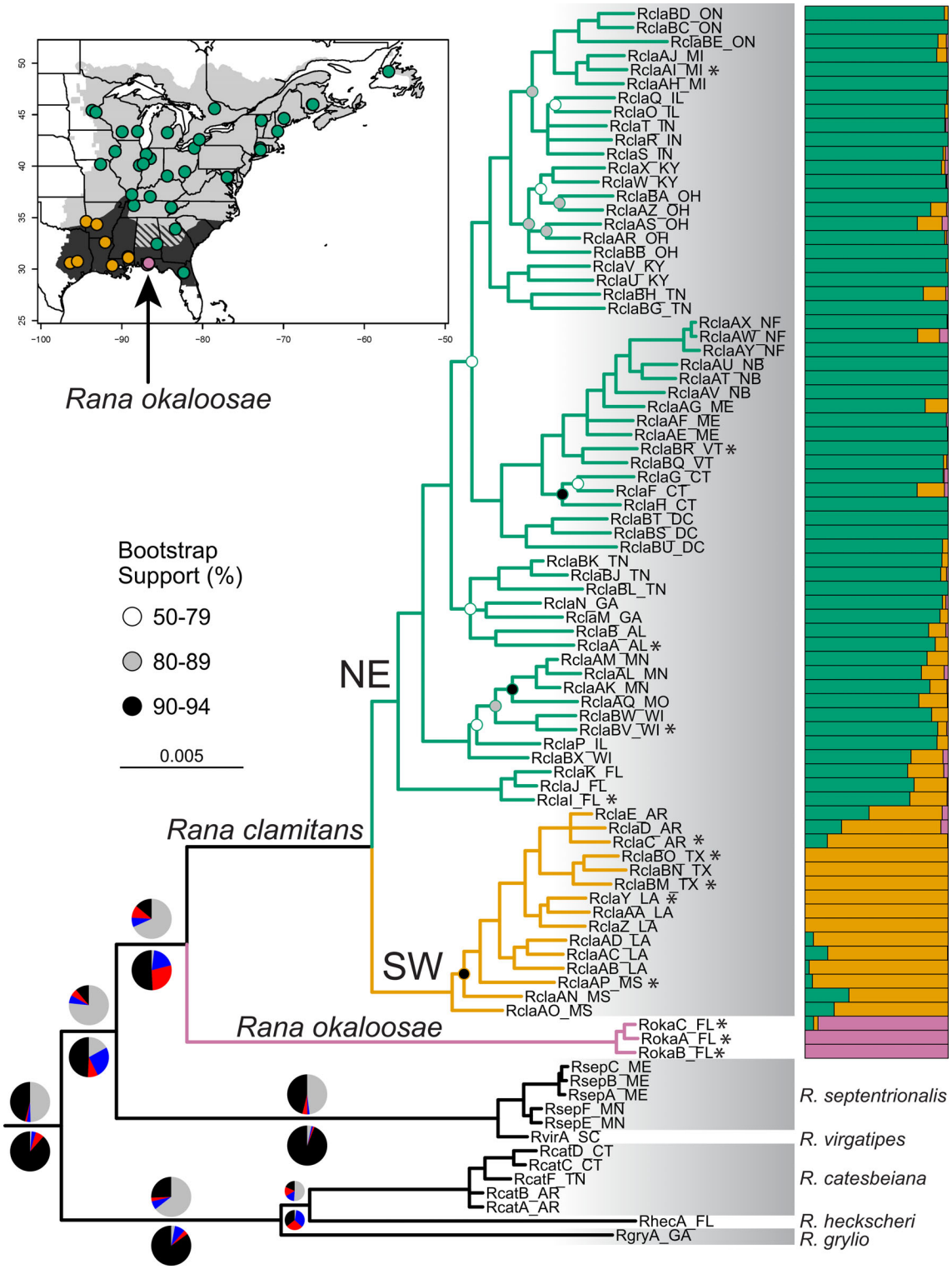

Fig. 2. Maximum likelihood ddRAD phylogeny of range-wide samples of *Rana clamitans* and all members of the *Aquarana* clade. Branch lengths are proportional to expected substitutions per site. Nodes with less than 50% bootstrap support are collapsed. Nodes with >95% bootstrap support are unlabeled. Inset is a map of our samples of *R. clamitans* against prior subspecies ranges (*Rana clamitans melanota* in gray, *R. c. clamitans* in black, hashed area represents the range of intergradation). *Rana clamitans* is comprised of two primary clades, a southwestern clade (orange) and a widespread northern clade (green). Bar plot adjacent to *R. clamitans* displays individual ancestry coefficients estimated by LEA for K = 3. Pie charts on the phylogeny indicate concordance factor (CF) support. Pie charts above branches represent gene CFs, while pie charts below represent site CFs. Black indicates agreement with the displayed tree, red and blue indicate disagreement, and gray indicates no support or noise. Asterisks next to tips indicate samples used for species delimitation analyses. See Data Accessibility for tree file.

populations from Florida (Figs. 2, S1; see Data Accessibility). Relationships within these two clades are not consistent across analyses of different ddRAD datasets, and bootstrap support is low to moderate for many nodes (Fig. S1; see Data Accessibility). The species tree analysis resolves two main clades within *R. clamitans*, with the specimens from Florida and the northern populations placed in a single clade (Fig. S2; see Data Accessibility). Relationships among sample sites within these clades are poorly supported (Fig. S2; see Data Accessibility). The mtDNA phylogeny is not consistent with the ddRAD phylogenies (Fig. S3; see Data Accessibility). For example, specimens from Florida are not resolved as a clade in the mtDNA phylogeny, and populations from Louisiana and Arkansas are broadly polyphyletic (Fig. S3; see Data Accessibility).

***Population structure of* Rana clamitans.**—We observe an overall signal of IBD within *Rana clamitans* that is additionally layered with important population structure. All three IBD datasets (Table S4; see Data Accessibility) produced similar results; therefore, we focus on results for the dataset with <30% missing data. There is a moderately strong positive correlation between geographic distance and genetic distance ($R^2 = 0.22$, $P < 0.0001$, Fig. 3C, D). This same pattern is confirmed by a Mantel test, which produced a Pearson correlation coefficient of 0.47, significantly higher than the mean Pearson correlation coefficient for 10,000 permuted datasets ($P < 1.0e{-}4$). However, AICc model comparison and model averaging find that models containing clade comparisons improved IBD inferences. Specifically, our analysis found similar support for a model containing an interaction between geographic distance and clade comparisons (log likelihood = 3131.524, AICc –6248.9, delta AIC = 0, weight 0.584, $R^2 = 0.68$, $P < 2.2e{-}16$) and a model containing an additive relationship between geographic distance and clade comparisons (log likelihood = 3129.153, AICc –6248.2, delta AIC = 0.68, weight 0.416, $R^2 = 0.68$, $P < 2.2e{-}16$). Both models with geographic distance perform much better than univariate models or the null model (all delta AICc > 100, model weights 0.0).

The interaction here suggests a steeper IBD relationship for the southwest clade but similar slopes for the within-northeast comparison and between northeast-southwest contrast (Fig. 3C). A Tukey's *post hoc* test on the additive model (and assuming no slope differences among the three clade comparisons) finds that genetic divergence is more pronounced for among-clade comparisons than genetic divergence within either clade when controlling for geographic distance ($P < 0.0001$; Fig. 3D). IBD patterns do not vary among the northeastern and southwestern clade ($P = 0.37$). Because these two models contain similar model weights, have nearly indistinguishable AICc, and have identical R^2 values, the simplest model is the additive model. Further support for the additive model over the interactive model comes from the datasets with <40% and <50% missing data, which find lower support for the interaction model compared to the additive model (<40% missing data: interaction model delta AICc = 0.98, model weight = 0.38; <50% missing data: interaction model delta AICc = 2.27, model weight = 0.24).

Across all population structure analyses using LEA, K = 2 consistently has the lowest cross-entropy for *Rana clamitans* alone (Fig. S4; see Data Accessibility), which recapitulates the northeastern and southwestern clades. When including *R. okaloosae*, LEA consistently identifies the lowest cross-entropy score for K = 3 (Fig. S4; see Data Accessibility). Cross-validation plots for tess3R also display an elbow that would indicate an optimum K = 2 for *R. clamitans* alone and K = 3 for *R. clamitans* plus *R. okaloosae* (Fig. S4; see Data Accessibility). However, tess3R cross-validation scores generally decrease as K increases, potentially indicating IBD (François, 2016). Genetic clustering results are largely insensitive to missing data thresholds for K = 2 with *R. clamitans* only (Fig. S5; see Data Accessibility) and for K = 3 with *R. clamitans* and *R. okaloosae* (Fig. S6; see Data Accessibility). For LEA analyses, results are also not sensitive to the sNMF regularization parameter α (Fig. S7; see Data Accessibility). Therefore, we focus on results from LEA and tess3R analyses using the dataset with <40% missing genotypes per SNP and $\alpha = 10$ (LEA only). This dataset represents a compromise between removing SNPs with large proportions of missing data while still retaining several thousand informative markers.

For K = 2 with only samples of *R. clamitans*, LEA ancestry coefficient bar plots reveal a southwestern (SW) genetic cluster with ~100% ancestry assigned to samples from Texas (TX), Louisiana (LA), and Mississippi (MS; Fig. 4). Multiple individuals from LA, MS, and Arkansas (AR) have mixed genetic ancestry. There are many individuals with ~100% ancestry from the widespread northeastern clade, though southwest ancestry is present in many northeastern populations, especially those from Florida and the upper Midwest. Incorporation of geographical priors with tess3R reveals qualitatively similar patterns as the LEA analyses (Fig. 4). Generally, as geographic distance from TX increases, the percent of southwestern clade ancestry decreases for LA, MS, and AR samples (Fig. 4).

When *R. okaloosae* is included in the analysis, the optimal K = 3 shows a very small proportion of *R. okaloosae* ancestry in some samples of *R. clamitans*, though this result is inconsistent across datasets (Fig. S6; see Data Accessibility). LEA analyses identify one individual of *R. okaloosae* (RokaC_FL) with approximately 5–15% ancestry derived from *R. clamitans*. However, similar admixture is not detected by tess3R (Fig. S6; see Data Accessibility). PCA reveals that the admixed ancestry of this individual of *R. okaloosae* is likely an artifact of missing data. In the dataset where SNPs are required to have <30% missing genotypes across all samples, nearly 58% of genotypes are missing for the putatively admixed individual of *R. okaloosae*. PCA with this dataset shows the putatively admixed individual of *R. okaloosae* has an intermediate score along PC1, which separates *R. okaloosae* from *R. clamitans* (Fig. 5). In contrast, PCA using the SNP dataset with no missing data reveals that the putatively admixed individual of *R. okaloosae* has a nearly identical PC1 score as the other two individuals of *R. okaloosae* (Fig. 5).

For larger, non-optimal K values for *Rana clamitans*, tess3R and LEA identify additional population structure. At K = 3, both analyses identify a genetic cluster corresponding to far-northeastern populations in New England, New Brunswick, and Newfoundland (Fig. 4). This genetic cluster appears clinal and geographically dispersed, with far-northeastern ancestry decreasing with geographic distance from New Brunswick. Even some individuals of *R. clamitans* from as far southwest as Louisiana have a small proportion of far-

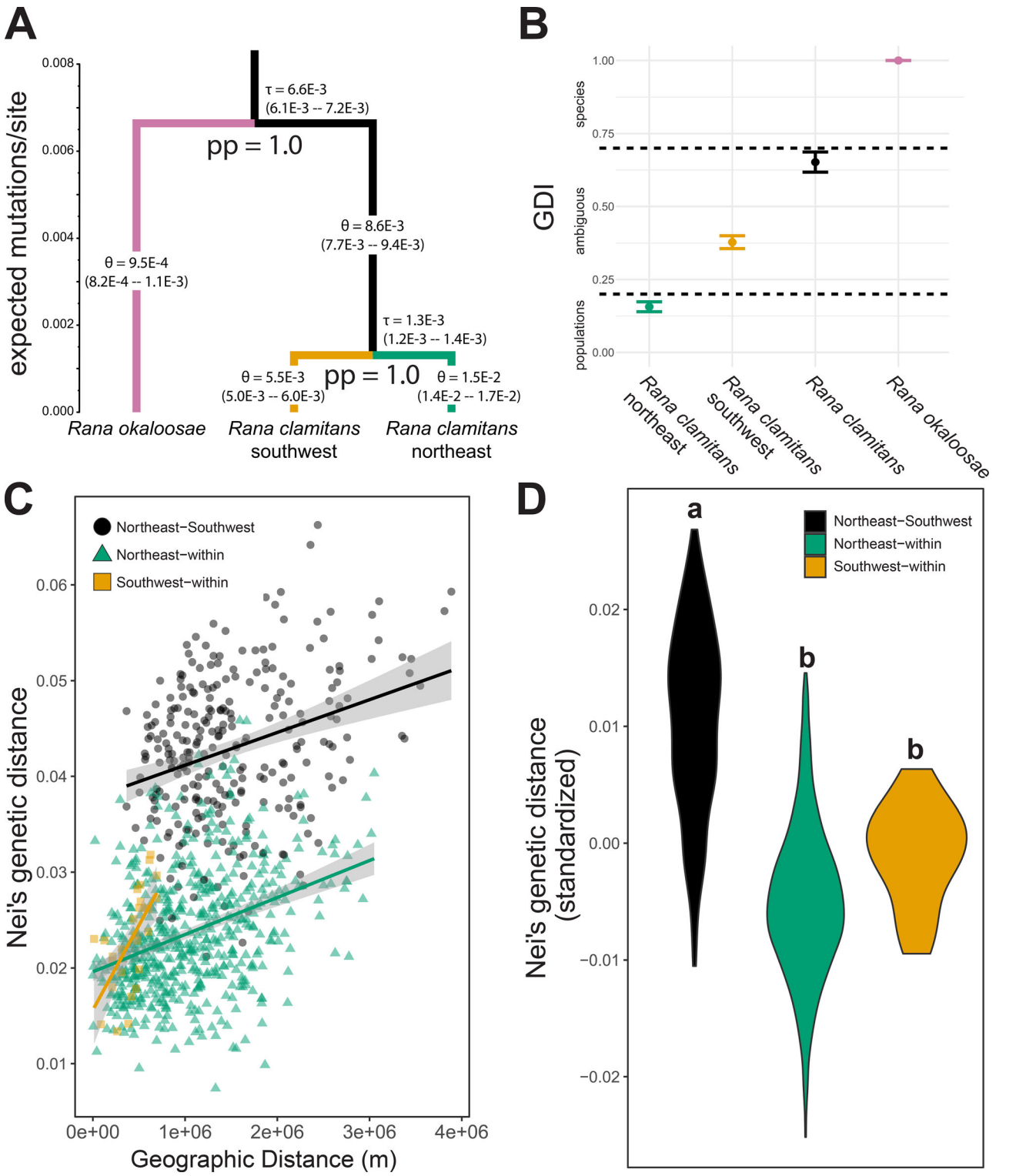

Fig. 3. (A–B) Species delimitation of *Rana clamitans* and *R. okaloosae*. (A) Guide species tree with theta estimates displayed along branches and tau estimates displayed above nodes, both given in units of expected mutations/site. BPP rjMCMC posterior probabilities below nodes indicate support for each speciation event. (B) Genealogical divergence index (GDI) means and 95% highest posterior density intervals for each tip in the guide tree, as well as for *R. clamitans* collapsed as a single species. (C–D) Genome-wide differentiation in *Rana clamitans* is largely a function of isolation-by-distance (IBD). (C) Linear model of genetic distance versus geographic distance, with separate regression lines for each comparison. (D) Violin plots of genetic distance standardized for geographic distance and IBD between northeast and southwest clades as well as within each clade. Letters above violin plots indicate Tukey HSD significant differences. Genetic distance values are the residuals of a regression of Nei's genetic distance against geographic distance. Between-clade comparisons show elevated genetic divergence compared to what would be predicted from geographic distance and IBD alone.

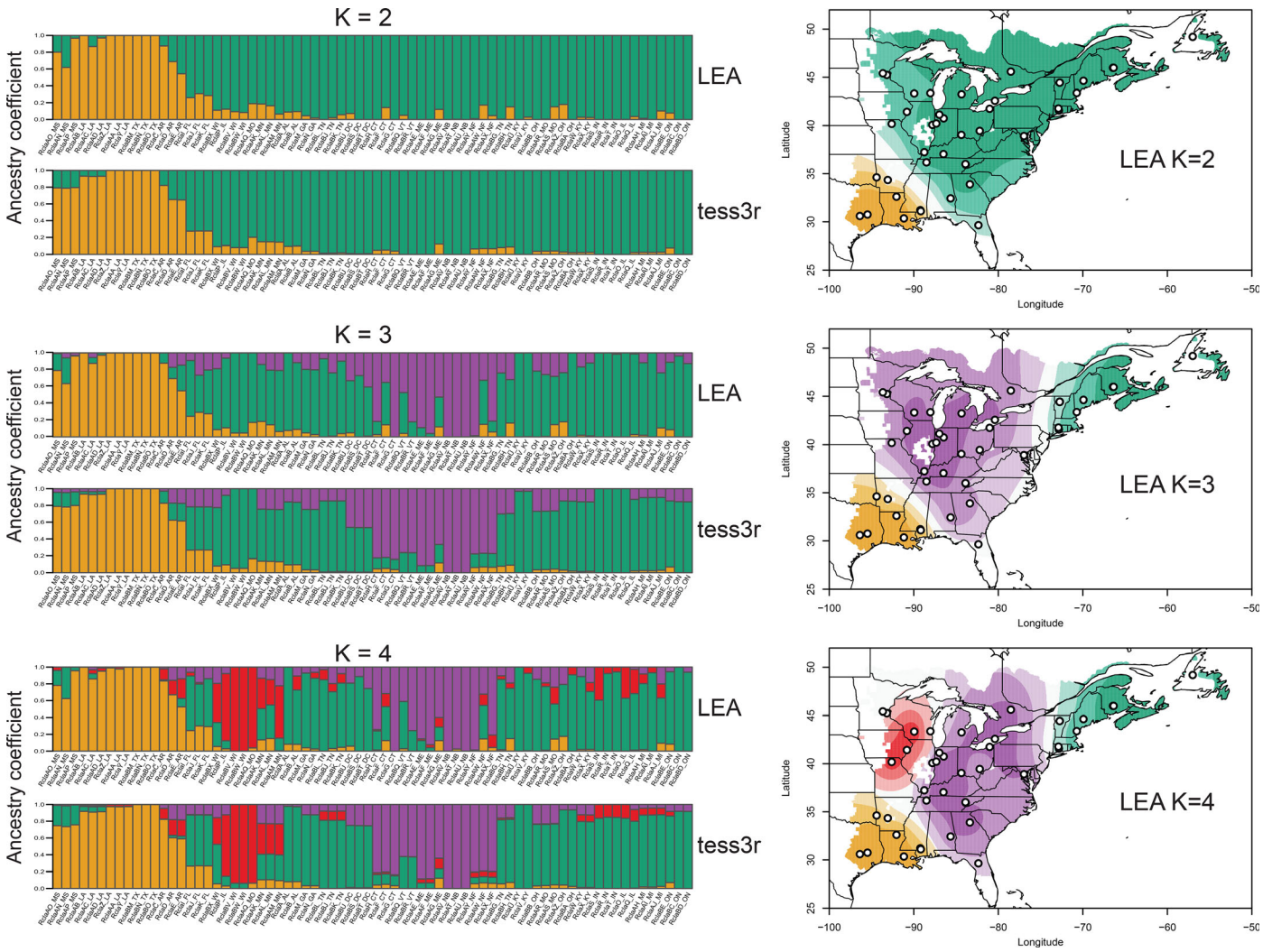

Fig. 4. Estimated ancestry coefficients of *Rana clamitans* for K = 2 (optimal), K = 3, and K = 4. Each bar plot column represents an individual, while colors correspond to estimated ancestry coefficients. LEA interpolated ancestry coefficients are displayed on the map, restricted to the known range of *R. clamitans*.

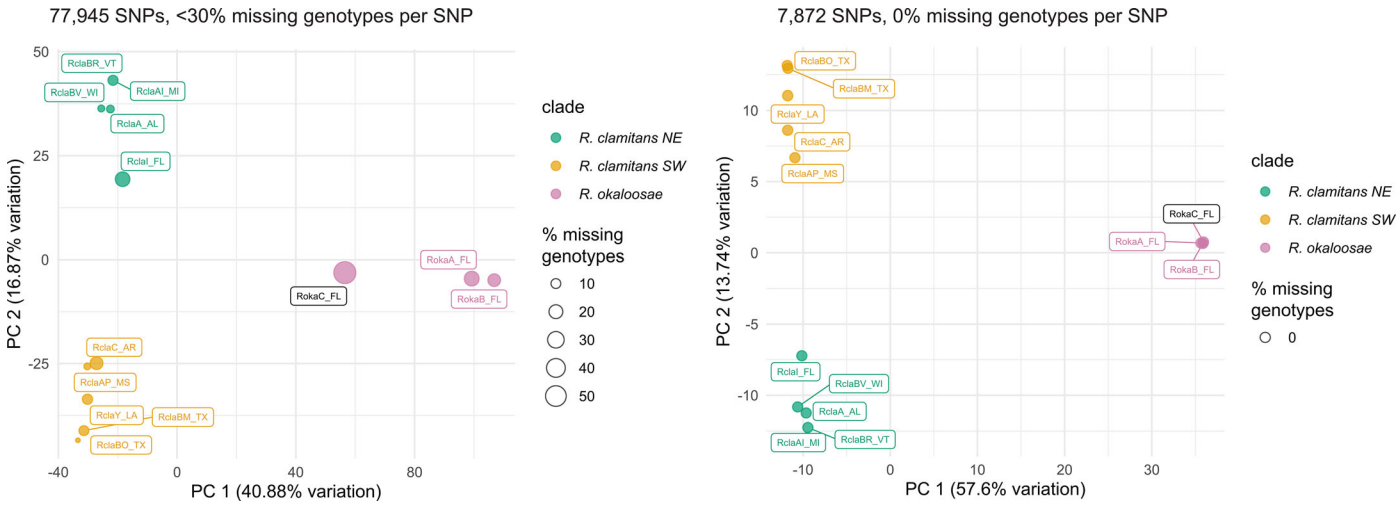

Fig. 5. PCA of ddRAD SNPs allowing a maximum of 30% missing genotypes per SNP or 0% missing genotypes per SNP. Size of points is proportional to the percentage of missing genotypes for that individual. The putative admixed individual (RokaC_FL) is highlighted in black. Colors correspond to *Rana okaloosae* plus the two major clades within *R. clamitans* (Fig. 2).

northeastern genetic ancestry. K = 4 ancestry coefficient bar plots identify a distinct genetic cluster in Wisconsin, Missouri, Illinois, and Minnesota (Fig. 4). Like the far-northeast genetic cluster, minor signals of ancestry from this K = 4 genetic cluster are also present in many other samples across the range of *R. clamitans*.

***Species delimitation of* Rana clamitans.**—We assembled a subset of 1,116 ddRAD loci shared by 13 individuals that encapsulate the breadth of genetic diversity in *Rana clamitans* and *R. okaloosae* (samples with asterisks, Fig. 2). The BPP rjMCMC species delimitation algorithm suggests that the northeast and southwest *R. clamitans* clades represent distinct species with posterior probability of 1.0 (Fig. 3A). However, the degree of genetic divergence for these species is very low, as demonstrated by the GDI. The mean GDI for the southwestern clade was 0.38 (95% credible interval = 0.36–0.40), while the mean GDI for the northeastern clade was 0.16 (95% credible interval = 0.14–0.17; Fig. 3B). The GDI estimate for the northeastern *R. clamitans* clade falls below the intraspecific GDI threshold of 0.2 suggested by Jackson et al. (2017). By contrast, the mean GDI estimate for the entire lineage of *R. clamitans* is much higher, with a mean value of 0.65 (95% credible interval = 0.62–0.68), falling just short of the Jackson et al. (2017) interspecific GDI threshold of 0.7 (Fig. 3B). Therefore, despite the BPP results, GDI estimates suggest that *R. clamitans* comprises a single species. In addition, *R. okaloosae* is strongly supported as a distinct species by both the BPP and GDI species delimitation approaches (Fig. 3).

***Phylogeny of* Aquarana.**—Concatenated ML ddRAD phylogenies all provide strong support for the same phylogenetic relationships among species of *Aquarana* (Figs. 1, 2, S1; see Data Accessibility). Of note, *Rana clamitans* and *R. okaloosae* are resolved as sister species. The ddRAD phylogeny contrasts with conclusions from previous mtDNA analyses (Austin et al., 2003) and our own analysis of mtDNA (Figs. 1, S3; see Data Accessibility), which resolve *R. okaloosae* nested within *R. clamitans*. The ML ddRAD phylogeny has few similarities with previously hypothesized evolutionary relationships of *Aquarana*, aside from a sister relationship between *R. catesbeiana* and *R. heckscheri* (Fig. 1).

Compared to the concatenated ML phylogeny, species tree analyses infer a nearly identical topology of *Aquarana* except for one node. In the species tree analyses, *Rana okaloosae* and *R. clamitans* are weakly supported as the sister lineage of a clade containing *R. grylio*, *R. heckscheri*, and *R. catesbeiana*. In contrast, the concatenated ML phylogeny resolves *R. okaloosae* and *R. clamitans* as the sister lineage of a clade containing *R. virgatipes* and *R. septentrionalis* with 100% bootstrap support (Fig. 2). Gene concordance factors indicate that a larger proportion of the ddRAD loci support the concatenated ML topology (11.1%) compared to the species tree topology (6.5%). Interestingly, site concordance factors show a much larger proportion of SNPs support the ML topology (49.2%) versus the species tree topology (25.6%). Thus, the ML topology (Fig. 1, 2) reflects our best hypothesis of the evolutionary history of the *Aquarana* clade.

The only other poorly supported node in the species tree is the clade containing *Rana heckscheri* and *R. catesbeiana* with bootstrap of 71% (Fig. 1). This node was also resolved in the ML concatenated phylogeny, but with 100% bootstrap support (Figs. 1, 2). Compared to the other contentious relationships among species of *Aquarana* discussed above, concordance factors are far less decisive. 18.4% of gene trees support a clade containing *R. heckscheri* and *R. catesbeiana*, while 15.2% support *R. grylio* and *R. catesbeiana* as sister species and 15.5% support *R. grylio* and *R. heckscheri* as sister species. Site concordance factors are also nearly evenly split between the three topologies (36% *R. heckscheri* + *R. catesbeiana*, 28% *R. grylio* + *R. catesbeiana*, and 34% *R. grylio* + *R. heckscheri*). Although the ddRAD species tree and ML concatenated phylogeny infer a sister relationship between *R. heckscheri* and *R. catesbeiana* as previously hypothesized (Hillis and Wilcox, 2005; Yuan et al., 2016), this specific resolution is characterized by phylogenetic discordance among the ddRAD loci.

TreeMix finds inconsistent support for gene flow among species of *Aquarana* and between populations of *R. clamitans*. In the phylogenetically broadest TreeMix analyses, approximately 75–77% of SNP subsets identify a strong signal of gene flow from northeast *R. clamitans* into *R. septentrionalis* with 16% mean migration weight (the percent ancestry derived from gene flow). 22–26% of SNP subsets identify similarly strong gene flow from southwest *R. clamitans* into *R. septentrionalis* (Fig. S9; see Data Accessibility). There are also weaker signals of gene flow from *R. catesbeiana* into both *R. clamitans* clades (mean migration weight = 2–3%) and from *R. catesbeiana* into *R. okaloosae* (mean migration weight = 0.002%; Fig. S9; see Data Accessibility). When analyzing only *R. clamitans* and *R. okaloosae*, TreeMix results differ substantially depending on whether the ML phylogeny or genetic clustering results are used to partition individuals into populations. TreeMix analyses with population assignment guided by the phylogeny find gene flow from the northeastern clade of *R. clamitans* into *R. okaloosae* supported by only 27–28% of SNP subsets. In contrast, when LEA genetic clustering guides population assignment, 71–80% of SNP subsets support gene flow from the far-northeastern genetic cluster of *R. clamitans* into *R. okaloosae* (Fig. S9; see Data Accessibility). Migration weights are generally larger in the analyses of only *R. clamitans* and *R. okaloosae* compared to the broader analyses of *Aquarana*.

DISCUSSION

Despite being one of the most locally abundant and widespread amphibians in North America, the evolutionary history of the Green Frog (*Rana clamitans*) has remained markedly unresolved for over a century. This includes delimitation of putative subspecies and evolutionary relationships with other species of *Aquarana* (Rhoads, 1895; Mecham, 1954; Austin et al., 2003; Austin and Zamudio, 2008; Hillis and Wilcox, 2005; Yuan et al., 2016). By analyzing hundreds of thousands of genomic markers, we illuminate the complex phylogeography of *R. clamitans*, offer a richer interpretation of subspecies, clarify possible hybridization with Vulnerable *R. okaloosae*, and provide greater resolution of the phylogeny of *Aquarana*.

Rana clamitans *is characterized by isolation-by-distance.*— Across the range of *Rana clamitans*, we find that the species is characterized largely by IBD, but with population structure defined by a large northeastern clade and comparatively smaller southwestern clade. Our conclusions contrast with

those of a previous molecular study that described more complex and discrete population structure (Austin and Zamudio, 2008). Earlier studies were likely misled by sparser genetic and geographic sampling that did not capture the overall IBD pattern. Despite spanning several well-known biogeographic barriers including the Appalachian Mountains, the Atlantic Fall Line, and the Alabama-Appalachian suture zone (Hoffman and Blouin, 2007; Gamble et al., 2008; Zeisset and Beebee, 2008; Rissler and Smith, 2010; Newman and Rissler, 2011), we find that genomic differentiation in *R. clamitans* is largely consistent with IBD (Fig. 3C, D). The primary axis of genetic variation within *R. clamitans* follows a broad cline from southwest to northeast. This pattern suggests that the relative vagility and generalist natural history of *R. clamitans* facilitates gene flow across complex biogeographic barriers. Indeed, Spring Peeper, *Pseudacris crucifer*, a similarly widespread and common frog species in eastern North America, also exhibits similar broad geographic patterns of gene flow, highlighting how vagile species can have limited population structure even at broad geographic scales (Cairns et al., 2021).

Our phylogenetic analyses further illuminate the genetic structure of populations across the Mississippi River, an area of the range undersampled in previous studies. We found consistent evidence for a phylogenetic split in *Rana clamitans* between a southwestern clade and a widespread northeastern clade (Fig. 2). This population structure explains the intraspecific diversity of *R. clamitans* additively in conjunction with IBD (Fig. 3C, D). The northeastern clade includes populations ranging across the Eastern Seaboard from Florida into New Brunswick and an invasive part of the species' range in Newfoundland and as far west as Minnesota toward the middle of the continent. Our results are incongruent with the only previous phylogeographic study of *R. clamitans*, which identified a Coastal Plain and Appalachia (CPA) clade that overlapped with a second "widespread clade" (Austin and Zamudio, 2008). We do not resolve these distinct sub-clades within the northeastern clade (Fig. S1; see Data Accessibility), and even with larger, non-optimal K values, none of the genetic clusters correspond to the previously described CPA clade (Fig. 4). The larger, non-optimal K values for tess3r and LEA generally recapitulate clinal patterns at smaller geographic scales, indicating that much of the genetic variation within *R. clamitans* is likely the primary result of IBD (Fig. 4).

One notable finding at larger K values is the existence of a northwestern genetic cluster with 100% ancestry assignment for two samples from southwestern Wisconsin (Fig. 4, red genetic cluster). Southwest Wisconsin is part of the Driftless Region, a high-latitude area that remained unglaciated during the Pleistocene (Holliday et al., 2002) and has been implicated as a refugium for many other species (Peck and Christiansen, 1990; Rowe et al., 2004; Li et al., 2013), including at least one other ranid frog (Lee-Yaw et al., 2008). Despite the genetic distinctiveness of the Wisconsin individuals of *R. clamitans*, these individuals are phylogenetically nested within a larger northwestern group spanning Illinois, Wisconsin, Missouri, and Minnesota (Fig. 2). Indeed, the northwestern genetic cluster is not restricted to the Driftless Region, as individuals from Illinois, Minnesota, and Missouri also contain substantial ancestry from this genetic cluster (Fig. 4). If the Driftless Region was a glacial refugium for *R. clamitans*, there has likely been post-Pleistocene secondary contact and gene flow as the species dispersed from more southerly refugia. This is supported by the presence of minor northwestern genetic ancestry across the range of *R. clamitans*. Our findings encourage further work on the role of the Driftless Region on population structure and gene flow in *R. clamitans* and other amphibians.

The southwestern *Rana clamitans* clade is concentrated in a relatively small region of the species' range, primarily in Texas and Louisiana. The southwestern clade generally follows the pattern recognized in multiple species with a break around the Mississippi River (Soltis et al., 2006; Myers et al., 2020). However, in *R. clamitans* the genetic variation is clinal across Louisiana and Arkansas rather than a sharp break at the Mississippi River. The southwestern clade identified here is similar to the Louisiana lineage identified by Austin and Zamudio (2008); however, sparse sampling in the southwestern portion of the species' range limited their study. A similar southwestern clade in widespread Spring Peeper frogs (*Pseudacris crucifer*) suggests shared biogeographic processes shaped the evolutionary history of both species and perhaps other (semi)aquatic species (Cairns et al., 2021).

Unlike other species with similar southern ranges, we did not observe distinct population genetic structure in Florida *Rana clamitans* (Solits et al., 2006; Austin and Zamudio, 2008). Our phylogenetic results suggest that Florida *R. clamitans* are sister to all other populations of the northeastern clade (Fig. 2), perhaps representing evidence for a Pleistocene refugium during glacial range reductions prior to subsequent range expansions northward (Newman and Rissler, 2011). However, this Florida clade was not identified as a distinct group in any genetic clustering analyses (Fig. 4). Prior inferences based on mtDNA suggested Florida populations of *R. clamitans* contain haplotypes from both Austin and Zamudio's (2008) CPA and other geographically widespread clades. In sum, prior work and our extensive genomic analyses here provide evidence for a large role for IBD rather than discrete geographic lineages, with a porous phylogeographic break distinguishing populations in Texas, Louisiana, Mississippi, and Arkansas versus populations from the rest of the geographic range.

Spurious species delimitation.—We find *R. clamitans* to be another case where coalescent species delimitation methods may misidentify population structure (in this case, clinal genetic variation) as distinct species (e.g., Chambers and Hillis, 2020; Mason et al., 2020). The standard rjMCMC BPP species delimitation methodology, which is widely used but also contested (Jackson et al., 2017; Sukumaran and Knowles, 2017; Leaché et al., 2019), suggests genomic differentiation within *R. clamitans* is strong enough to support the southwestern and northeastern clades as separate species (Fig. 3C). However, the low GDI of these two clades contradicts the BPP delimitation, instead suggesting that *R. clamitans* is only a single species (Fig. 3D). Furthermore, population genomic analyses reveal clinal variation within *R. clamitans* that undergirds broad genetic breaks among populations and clades. Therefore, in aggregate, our work does not support recognition of multiple independent evolutionary lineages within *R. clamitans* and casts additional doubt on species delimitations that rely solely on methods like BPP.

***Subspecies of* Rana clamitans *are of little utility*.**—Although species are generally agreed to be a fundamental as well as utilitarian level of biological organization, the utility of subspecies as a taxonomic unit is particularly controversial (Mayr, 1982; Burbrink et al., 2000; Patten and Remsen, 2017; de Queiroz, 2020). Our analysis recapitulates some aspects of Mecham's (1954) morphology-based subspecies delimitations with respect to identifying meaningful evolutionary lineages, but not along the original geographic boundaries. Mecham (1954) suggested two subspecies of *R. clamitans* occurred with a boundary along the coastal plain based on limited morphology. However, these subspecies are no longer recognized by herpetologists (Crother, 2017; Stebbins and McGinnis, 2018). The southwestern clade we identified represents the western half of Mecham's (1954) delimitation of *R. c. clamitans*. Interestingly, the samples Mecham (1954) originally assigned to *R. c. clamitans* largely came from Texas, Arkansas, Louisiana, and Mississippi with relatively sparse sampling from Alabama eastward in the rest of the former subspecies' range. However, the type locality for *R. c. clamitans* is inherently the type locality for *R. clamitans* in South Carolina, far from the boundaries of our southwestern clade (Stewart, 1984). If the southwestern clade of *R. clamitans* were to be defined as a subspecies, a new epithet would be needed as *R. c. clamitans* is not applicable.

The phylogenetic break between the widespread northeastern and southwestern clades lends support for the delimitation of two distinct subspecies of *Rana clamitans*. However, population structure analyses reveal that these clades lack a sharp genetic break and exhibit admixture indicative of clinal IBD. The pattern of IBD we identified does suggest that between-clade genetic variation more richly describes genetic variation in conjunction with geographic distance. Additionally, designating these two clades as subspecies would be inconsistent with the morphological variation described by Mecham (1954), as his phenotypic subspecies do not follow the geographic boundaries of the northeast and southwest clades. The lack of associated phenotypic and genetic variation means that delimiting the northeastern and southwestern clades as subspecies would have no practical utility for herpetologists or evolutionary biologists. However, it is possible that these clades are phenotypically divergent in ways not detected by Mecham's (1954) qualitative assessment of anatomical variation and coloration. We recommend further sampling across the putative cline between the two major clades of *R. clamitans* in addition to thorough morphological and ecological analyses. Advances in quantitative color, acoustic, and morphometric analyses are poised to thoroughly characterize phenotypic differences among clades of *R. clamitans*.

***Limited evidence of hybridization between* Rana clamitans *and* Rana okaloosae.**—Our analyses using ddRAD markers resolve *R. clamitans* and *R. okaloosae* as reciprocally monophyletic sister species. This contrasts with mtDNA phylogenetic analyses (Austin et al., 2003; Austin and Zamudio, 2008; Fig. S3; see Data Accessibility), which find *R. okaloosae* phylogenetically nested within *R. clamitans*. Such phylogenetic discordance between nuclear markers and mtDNA is common and often the result of introgressive hybridization (Chan and Levin, 2005; Bossu and Near, 2009; Denton et al., 2014; Firneno et al., 2020). In addition, LEA genetic clustering analyses identify one sample of *R. okaloosae* (RokaC_FL) with a small proportion of *R. clamitans* ancestry, suggestive of hybridization and introgression between Vulnerable *R. okaloosae* and the widespread *R. clamitans*. However, PCA reveals that substantial missing data at the individual level produces misleading inferences of admixture and hybridization (Fig. 5). The apparently admixed ancestry of RokaC_FL is likely the product of a large proportion of missing genotypes rather than hybridization and gene flow (Fig. S8; see Data Accessibility). Our work should serve as a cautionary tale when investigating admixture and gene flow using unevenly sparse datasets, as is often the case with RAD sequencing.

TreeMix analyses provide mixed support for gene flow between *R. clamitans* and *R. okaloosae*. When analyzing the broader *Aquarana* clade, no gene flow is detected between *R. clamitans* and *R. okaloosae* (Fig. S9; see Data Accessibility). However, when analyzing only *R. clamitans* and *R. okaloosae*, TreeMix identifies substantial gene flow from northeastern *R. clamitans* into *R. okaloosae* (Fig. S9; see Data Accessibility). The strength of this inference is dependent on the method used to assign individuals of *R. clamitans* to populations. Only 27–28% of SNP subsets identify gene flow (mean migration weight = 33%) from the northeastern *R. clamitans* clade into *R. okaloosae* when phylogeny is used for population assignment, whereas 71–80% of SNP subsets identify gene flow (mean migration weight = 54%) when population assignment corresponds to LEA K = 4 genetic clusters (Fig. S9; see Data Accessibility). Both sets of analyses also identify a weak signal of gene flow (mean migration weight = 2%) from *R. okaloosae* into the northeastern *R. clamitans* clade or the far-northeastern genetic cluster of *R. clamitans* (Fig. S9; see Data Accessibility).

Although TreeMix analyses of the LEA K = 4 genetic clusters strongly support gene flow from *R. clamitans* into *R. okaloosae*, the inferred geographic origin of this gene flow seems implausible. The strongest signal of gene flow into *R. clamitans* originates from the far-northeastern genetic cluster of *R. clamitans* in New England, New Brunswick, and Nova Scotia (Rcla_NE, Fig. S9; see Data Accessibility). These populations are geographically distant from *R. okaloosae*, which is found only in Florida. Only 18% of SNP subsets identify gene flow from the genetic cluster of *R. clamitans* that is geographically proximate to *R. okaloosae* (Rcla_mid, Fig. S9; see Data Accessibility). Our limited sampling prevents detailed analysis of hybridization and introgression at fine spatial scales. It remains unclear whether mtDNA capture between *R. okaloosae* and *R. clamitans* is the result of historical or contemporary hybridization. Dense sampling of *R. clamitans* surrounding the range of *R. okaloosae* will help elucidate the history of gene flow (Austin et al., 2011). Such work should be a priority for conservation of the Vulnerable *R. okaloosae*, as hybridization can lead to invasive gene flow (Mallet, 2005), ultimately eroding the genetic integrity of this unique lineage.

***Phylogeny of* Aquarana.**—Despite strong support for monophyly, our understanding of phylogenetic relationships within the *Aquarana* clade has remained markedly unresolved (Austin et al., 2003; Austin and Zamudio, 2008; Hillis and Wilcox, 2005; Yuan et al., 2016). Analysis of ddRAD data adds clarity to the evolutionary relationships among *Aquarana* and strongly supports resolution of three clades in *Aquarana*: the clade comprising *Rana clamitans* and *R.*

okaloosae; a clade containing *R. septentrionalis* and *R. virgatipes*; and a clade comprising *R. grylio*, *R. catesbeiana*, and *R. heckscheri*. Our results corroborate prior studies that identify close evolutionary relationships between *R. okaloosae* and *R. clamitans* (Austin et al., 2003; Hillis and Wilcox, 2005; Austin and Zamudio, 2008; Yuan et al., 2016), as well as *R. catesbeiana* and *R. heckscheri* as sister species (Hillis and Wilcox, 2005; Yuan et al., 2016). The short branch length leading to *R. catesbeiana* and *R. heckscheri* may explain the significant gene tree discordance and moderate species tree support at this node (Figs. 1, 2). While prior studies have inconsistently resolved the placement of *R. septentrionalis* and *R. virgatipes* within *Aquarana*, analyses of ddRAD markers provide strong support for their phylogenetic resolution as sister species.

Although our phylogenetic analyses find strong support for three subclades of *Aquarana*, there is inconsistent resolution of their relationships (Fig. 2). Concordance factors indicate that gene tree discordance is the source of this phylogenetic incongruence (Fig. 2). The ML concatenated topology is supported by the largest proportion of gene trees and SNPs, and thus is likely the best representation of the phylogeny of *Aquarana*. Reticulate evolution could provide an explanation for the phylogenetic discordance at this node, as TreeMix consistently identifies gene flow between *R. clamitans* and *R. septentrionalis* (Fig. S9; see Data Accessibility), which could confound phylogenetic inference. Alternatively, phylogenetic discordance among the subclades of *Aquarana* could be the result of substitutional saturation or non-orthologous RAD loci. However, these explanations are unlikely for two reasons. First, concordance factors indicate decisive topological support for the crown of *Aquarana*, the deepest phylogenetic node in our analyses (Fig. 2). Additionally, *Aquarana* has a crown age of approximately 15 million years (Yuan et al., 2016), well within estimated phylogenetic informativeness boundaries for RAD loci (Near et al., 2018). Future work with expanded taxonomic and genomic sampling can help further clarify these deeper nodes within *Aquarana* and identify potential causes of gene tree discordance. Overall, our phylogenomic analysis improves resolution of the phylogenetic relationships of *Aquarana* and provides an evolutionary framework for comparative ecological and evolutionary studies on ranid frogs in eastern and central North America.

Conclusions.—Analysis of genome-wide ddRAD markers identifies broad clinal genetic variation within *Rana clamitans* due to isolation-by-distance, with remarkably little additional population structure other than broad northeastern and southwestern clades. Despite contradictory results from genetic species delimitation analyses, our results support continued recognition of *R. clamitans* as a single species. Indeed, our work highlights the pitfalls of delimiting species using only a single methodology. Additionally, we find that the subspecies concept has little use in *R. clamitans* given the clinal nature of genetic variation, which does not correspond to poor existing descriptions of phenotypic variation. Unlike previous studies, we find that *R. clamitans* and *R. okaloosae* are phylogenetically distinct species with no conclusive evidence of hybridization. Finally, our results add clarity to phylogenetic relationships in the *Aquarana* clade, although there is still some uncertainty regarding one of the deeper nodes. Our work here provides an evolutionary framework to propel future research using *R. clamitans* and the *Aquarana* clade to address important ecological and evolutionary questions.

DATA ACCESSIBILITY

Raw sequence data have been deposited at NCBI GenBank or SRA (Table S1). New sequences were deposited at NCBI GenBank (accession numbers OP414959–OP415048) or SRA (accession numbers SRS15085385–SRS15085474). Alignments, VCF files, treefiles, and scripts are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.95x69p8mj. Supplemental material is available at https://www.ichthyologyandherpetology.org/h2021129. Unless an alternative copyright or statement noting that a figure is reprinted from a previous source is noted in a figure caption, the published images and illustrations in this article are licensed by the American Society of Ichthyologists and Herpetologists for use if the use includes a citation to the original source (American Society of Ichthyologists and Herpetologists, the DOI of the *Ichthyology & Herpetology* article, and any individual image credits listed in the figure caption) in accordance with the Creative Commons Attribution CC BY License.

ACKNOWLEDGMENTS

We thank Angus Mossman, Oliver Orr, Cathy Newman, Adam Clause, Becca Cozad, Chris Smith, Phil Pearson, and Kelly Irwin for assisting in collecting frog specimens and James Austin for providing samples from multiple species of *Aquarana*. We are grateful to the collections that provided tissue samples for this project, including the Field Museum of Natural History, Illinois Natural History Survey, New Brunswick Museum, and Yale Peabody Museum of Natural History. We thank Alan Resetar and Trevor Persons for additional samples and specimens. DJM, GGM, and MRL designed the research. DJM, MRL, GWC, and TJN secured funding for the project. GWC obtained necessary collecting permits. DJM and MRL collected samples and genomic data. DJM, GGM, and MRL analyzed data. All authors contributed to writing the paper. This work was funded by the National Science Foundation Division of Environmental Biology (grant number 1701311) to MRL and the Yale Peabody Museum Ginsberg Fund. DJM was supported by the Yale Training Program in Genetics (National Institute of Health grant number T32 GM007499).

LITERATURE CITED

AmphibiaWeb. 2020. https://amphibiaweb.org. University of California, Berkeley, California (accessed 17 September 2020).

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc

Austin, J. D., T. A. Gorman, and D. Bishop. 2011. Assessing the fine-scale genetic structure and relatedness in the micro-endemic Florida bog frog. Conservation Genetics 12: 833–838.

Austin, J. D., S. C. Lougheed, P. E. Moler, and P. T. Boag. 2003. Phylogenetics, zoogeography, and the role of dispersal and vicariance in the evolution of the *Rana*

catesbeiana (Anura: Ranidae) species group. Biological Journal of the Linnean Society 80:601–624.

Austin, J. D., and K. R. Zamudio. 2008. Incongruence in the pattern and timing of intra-specific diversification in bronze and bullfrogs (Ranidae). Molecular Phylogenetics and Evolution 48:1041–1053.

Bartón, K. 2009. Mu-MIn: multi-model inference. R Package Version 0.12.2/r18. https://CRAN.R-project.org/package=MuMIn

Bossu, C. M., and T. J. Near. 2009. Gene trees reveal repeated instances of mitochondrial DNA introgression in orangethroat darters (Percidae: *Etheostoma*). Systematic Biology 58:114–129.

Burbrink, F. T., R. Lawson, and J. B. Slowinski. 2000. Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept. Evolution 54:2107–2118.

Burbrink, F. T., and S. Ruane. 2021. Contemporary philosophy and methods for studying speciation and delimiting species. Ichthyology & Herpetology 109:874–894.

Cairns, N. A., A. S. Cicchino, K. A. Stewart, J. D. Austin, and S. C. Lougheed. 2021. Cytonuclear discordance, reticulation and cryptic diversity in one of North America's most common frogs. Molecular Phylogenetics and Evolution 156:107042.

Caye, K., and O. Francois. 2016. tess3R: inference of spatial population genetic structure. R package version 1.1.0.

Chambers, E. A., and D. M. Hillis. 2020. The multispecies coalescent over-splits species in the case of geographically widespread taxa. Systematic Biology 69:184–193.

Chan, K. M. A., and S. A. Levin. 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution 59:720–729.

Chan, K. O., R. K. Abraham, M. B. Sanguila, and R. M. Brown. 2020. Over-splitting destabilizes the taxonomy of hylaranine frogs: a response to Chandromouli et al. (2020). Zootaxa 4877:598–600.

Crother, B. I. (Ed.). 2017. Scientific and Standard English Names of Amphibians and Reptiles of North America North of Mexico, with Comments Regarding Confidence in Our Understanding. SSAR Herpetological Circular 43.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, and G. McVean. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

de Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology 56:879–886.

de Queiroz, K. 2020. An updated concept of subspecies resolves a dispute about the taxonomy of incompletely separated lineages. Herpetological Review 51:459–461.

Denton, R. D., L. J. Kenyon, K. R. Greenwald, and H. Gibbs. 2014. Evolutionary basis of mitonuclear discordance between sister species of mole salamanders (*Ambystoma* sp.). Molecular Ecology 23:2811–2824.

Dincă, V., K. M. Lee, R. Vila, and M. Mutanen. 2019. The condundrum of species delimitation: a genomic perspective on a mitogenetically super-variable butterfly. Proceedings of the Royal Society B 286:20191311.

Dray, S., and A.-B. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software 22:1–20.

Dufresnes, C., G. Mazepa, N. Rodrigues, A. Brelsford, S. N. Litvinchuk, R. Sermier, G. Lavanchy, C. Betto-Colliard, O. Blaser, A. Borzée, E. Cavoto, G. Fabre, K. Ghali, C. Grossen . . . D. L. Jeffries. 2018. Genomic evidence for cryptic speciation in tree frogs from the Apennine Peninsula, with description of *Hyla perrini* sp. nov. Frontiers in Ecology and Evolution 6:144.

Eaton, D. A., and I. Overcast. 2020. ipyrad: interactive assembly and analysis of RADseq datasets. Bioinformatics 36:2592–2594.

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

Firneno, T. J., J. R. O'Neill, D. M. Portik, A. H. Emery, J. H. Townsend, and M. K. Fujita. 2020. Finding complexity in complexes: assessing the causes of mitonuclear discordance in a problematic species complex of Mesoamerican toads. Molecular Ecology 29:3543–3559.

Flouri, T., X. Jiao, B. Rannala, and Z. Yang. 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent. Molecular Biology and Evolution 35:2585–2593.

François, O. 2016. Running structure-like population genetic analyses with R. R Tutorials in Population Genetics, U. Grenoble-Alpes.

Frichot, E., and O. François. 2015. LEA: an R package for landscape and ecological association studies. Methods in Ecology and Evolution 6:925–929.

Frost, D. R. 2022. Amphibian Species of the World: an Online Reference. Version 6.1 (accessed January 21, 2022). Electronic Database accessible at https://amphibiansoftheworld.amnh.org/index.php. American Museum of Natural History, New York.

Frost, D. R., T. Grant, J. Faivovich, R. H. Bain, A. Haas, C. F. B. Haddad, R. O. de Sá, A. Channing, M. Wilkinson, S. C. Donnellan, C. J. Raxworthy, J. A. Campbell, B. L. Blotto, P. E. Moler . . . W. Wheeler. 2006. The amphibian tree of life. Bulletin of the American Museum of Natural History 279:1–370.

Frost, D. R., T. Grant, J. Faivovich, R. H. Bain, A. Haas, C. F. B. Haddad, R. O. de Sá, A. Channing, M. Wilkinson, S. C. Donnellan, C. J. Raxworthy, J. A. Campbell, B. L. Blotto, P. Moler . . . W. C. Wheeler. 2008. Is *The Amphibian Tree of Life* really fatally flawed? Cladistics 24:385–395.

Gamble, T., P. B. Berendzen, H. B. Shaffer, D. E. Starkey, and A. M. Simons. 2008. Species limits and phylogeography of North American cricket frogs (*Acris*: Hylidae). Molecular Phylogenetics and Evolution 48:112–125.

Goebel, A. M., J. M. Donnelly, and M. E. Atz. 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. Molecular Phylogenetics and Evolution 11:163–199.

Good, J. M., S. Hird, N. Reid, J. R. Demboski, S. J. Steppan, T. R. Martin-Nims, and J. Sullivan. 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. Molecular Ecology 17:1313–1327.

Gorman, T. A., D. C. Bishop, and C. A. Haas. 2009. Spatial interactions between two species of frogs: *Rana okaloosae* and *R. clamitans clamitans*. Copeia 2009:138–141.

Gorman, T. A., and C. A. Haas. 2011. Seasonal microhabitat selection and use of syntopic populations of *Lithobates okaloosae* and *Lithobates clamitans clamitans*. Journal of Herpetology 45:313–318.

Hijmans, R. J., E. Williams, and C. Vennes. 2016. Geosphere: spherical trigonometry. R package.

Hillis, D. M. 2019. Species delimitation in herpetology. Journal of Herpetology 53:3–12.

Hillis, D. M. 2020. The detection and naming of geographic variation within species. Herpetological Review 51:52–56.

Hillis, D. M., E. A. Chambers, and T. J. Devitt. 2021. Contemporary methods and evidence for species delimitation. Ichthyology & Herpetology 109:895–903.

Hillis, D. M., and S. K. Davis. 1986. Evolution of ribosomal DNA: fifty million years of recorded history in the frog genus *Rana*. Evolution 40:1275–1288.

Hillis, D. M., and T. P. Wilcox. 2005. Phylogeny of the New World true frogs (*Rana*). Molecular Phylogenetics and Evolution 34:299–314.

Hoang, D. T., O. Chernomor, A. Von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Molecular Biology and Evolution 35:518–522.

Hoffman, E. A., and M. S. Blouin. 2007. Evolutionary history of the northern leopard frog: reconstruction of phylogeny, phylogeography, and historical changes in population demography from mitochondrial DNA. Evolution 58:145–159.

Holliday, V. T., J. C. Knox, G. L. Running IV, R. D. Mandel, and C. R. Ferring. 2002. The central lowlands and Great Plains, p. 335–362. *In*: The Physical Geography of North America. A. Orme (ed.). Oxford University Press, New York.

Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. Biometrical Journal 50:346–363.

Jackson, N. D., B. C. Carstens, A. E. Morales, and B. C. O'Meara. 2017. Species delimitation with gene flow. Systematic Biology 66:799–812.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.

Kalyaanamoorthy, S., B. Q. Minh, T. K. Wong, A. von Haeseler, and L. S. Jermiin. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14:587–589.

Kassler, T. W., J. B. Koppelman, T. J. Near, C. B. Dillman, J. Levengood, D. L. Swofford, J. L. VanOrman, J. E. Claussen, and D. P. Philipp. 2002. Molecular and morphological analyses of the black basses: implications for taxonomy and conservation. American Fisheries Society Symposium 31:291–322.

Leaché, A. D., T. Zhu, B. Rannala, and Z. Yang. 2019. The spectre of too many species. Systematic Biology 68:168–181.

Lee-Yaw, J. A., J. T. Irwin, and D. M. Green. 2008. Postglacial range expansion from northern refugia by the wood frog, *Rana sylvatica*. Molecular Ecology 17:867–884.

Li, P., M. Li, Y. Shi, Y. Zhao, Y. Wan, C. Fu, and K. M. Cameron. 2013. Phylogeography of North American herbaceous *Smilax* (Smilacaceae): combined AFLP and cpDNA data support a northern refugium in the Driftless Area. American Journal of Botany 100:801–814.

MacGuigan, D. J., A. J. Geneva, and R. E. Glor. 2017. A genomic assessment of species boundaries and hybridization in a group of highly polymorphic anoles (*distichus* species complex). Ecology and Evolution 7:3657–3671.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in Ecology and Evolution 20:229–37.

Marshall, T. L., E. A. Chambers, M. V. Matz, and D. M. Hillis. 2021. How mitonuclear discordance and geographic variation have confounded species boundaries in a widely studied snake. Molecular Phylogenetics and Evolution 126: 107194.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10–12.

Mason, N. A., N. K. Fletcher, B. A. Gill, W. C. Funk, and K. R. Zamudio. 2020. Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance. Systematics and Biodiversity 18:269–280.

Mayr, E. 1982. Of what use are subspecies? Auk 99:593–595.

McCartney-Melstad, E., M. Gidis, and H. B. Shaffer. 2018. Population genomic data reveal extreme geographic subdivision and novel conservation actions for the declining foothill yellow-legged frog. Heredity 121:112–125.

Mecham, J. S. 1954. Geographic variation in the green frog *Rana clamitans* LaTreille. Texas Journal of Science 6:1–24.

Moler, P. E. 1985. A new species of frog (Ranidae: *Rana*) from northwestern Florida. Copeia 1985:579–583.

Moritz, C., C. J. Schneider, and D. B. Wake. 1992. Evolutionary relationships within *Ensatina eschscholtzii* complex confirm the ring species interpretation. Systematic Biology 41:273–291.

Myers, E. A., A. D. McKelvy, and F. T. Burbrink. 2020. Biogeographic barriers, Pleistocene refugia, and climatic gradients in the southeastern Nearctic drive diversification in cornsnakes (*Pantherophis guttatus* complex). Molecular Ecology 29:797–811.

Near, T. J., C. M. Bossu, G. S. Bradburd, R. L. Carlson, R. C. Harrington, P. R. Hollingsworth Jr., B. P. Keck, and D. A. Etnier. 2011. Phylogeny and temporal diversification of darters (Percidae: Etheostomatinae). Systematic Biology 60: 565–595.

Near, T. J., D. J. MacGuigan, E. Parker, C. D. Struthers, C. D. Jones, and A. Dornburg. 2018. Phylogenetic analysis of Antarctic notothenioids illuminates the utility of RADseq for resolving Cenozoic adaptive radiations. Molecular Phylogenetics and Evolution 129:268–279.

Newman, C. E., and L. J. Rissler. 2011. Phylogeographic analyses of the southern leopard frog: the impact of geography and climate on the distribution of genetic lineages vs. subspecies. Molecular Ecology 20:5295–5312.

Nguyen, L., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32:268–274.

Niemiller, M. L., T. J. Near, and B. M. Fitzpatrick. 2011. Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). Evolution 66:846–866.

Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.

Patten, M. A., and J. V. Remsen Jr. 2017. Complementary roles of phenotype and genotype in subspecies delimitation. Journal of Heredity 108:462–464.

Pauly, G. B., D. M. Hillis, and D. C. Cannatella. 2009. Taxonomic freedom and the role of official lists of species names. Herpetologica 65:115–128.

Peck, S. B., and K. Christiansen. 1990. Evolution and zoogeography of the invertebrate cave faunas of the Driftless Area of the Upper Mississippi River Valley of Iowa, Minnesota, Wisconsin, and Illinois, USA. Canadian Journal of Zoology 68:73–88.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135.

Pfeifer, B., U. Wittelsbuerger, S. E. Ramos-Onsins, and M. J. Lercher. 2014. PopGenome: an efficient Swiss army knife for population genomic analyses in R. Molecular Biology and Evolution 31:1929–1936.

Pickrell, J., and J. Pritchard. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genetics 8:e1002967.

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S. Wu, Y. Manes, S. Dreisigacker, J. Crossa, H. Sanchez-Villeda, M. Sorrells, and J. L. Jannink. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5:103–113.

Priestley, A. S., T. A. Gorman, and C. A. Haas. 2010. Comparative morphology and identification of Florida bog frog and bronze frog tadpoles. Florida Scientist 73:20–26.

Reyes-Velasco, J., J. D. Manthey, Y. Bourgeois, X. Freilich, and S. Boissinot. 2018. Revisiting the phylogeography, demography and taxonomy of the frog genus *Ptychadena* in the Ethiopian highlands with the use of genome-wide SNP data. PLoS ONE 13:e0190440.

Rhoads, S. N. 1895. Contributions to the zoology of Tennessee. No. 1, Reptiles and amphibians. Proceedings of the Academy of Natural Sciences of Philadelphia 47: 376–407.

Rissler, L. J., and W. H. Smith. 2010. Mapping amphibian contact zones and phylogeographical break hotspots across the United States. Molecular Ecology 19:5404–5416.

Rowe, K. C., E. J. Heske, P. W. Brown, and K. N. Paige. 2004. Surviving the ice: northern refugia and postglacial colonization. Proceedings of the National Academy of Sciences of the United States of America 101:10355–10359.

Shirose, L., and R. J. Brooks. 1995. Growth rate and age at maturity in syntopic populations of *Rana clamitans* and *Rana septentrionalis* in central Ontario. Canadian Journal of Zoology 73:1468–1473.

Soltis, D. E., A. B. Morris, J. S. McLachlan, P. S. Manos, and P. S. Soltis. 2006. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology 15: 4261–4293.

Stebbins, R. C., and S. M. McGinnis. 2018. Peterson Field Guide to Western Reptiles & Amphibians. Houghton Mifflin Harcourt, Boston.

Stewart, M. M. 1984. Redescription of the type of *Rana clamitans*. Copeia 1984:210–213.

Sukumaran, J., and L. L. Knowles. 2017. Multispecies coalescent delimits structure, not species. Proceedings of the National Academy of Sciences of the United States of America 114:1607–1612.

Sun, Y.-B., Z.-J. Xiong, X.-Y. Xiang, S.-P. Liu, W.-W. Zhou, X.-L. Tu, L. Zhong, L. Wang, D.-D. Wu, B.-L. Zhang, C.-L. Zhu, M.-M. Yang, H.-M. Chen, F. Li . . . Y.-P. Zhang. 2015. Whole-genome sequence of the Tibetan frog *Nanorana parkeri* and the comparative evolution of tetrapod genomes. Proceedings of the National Academy of Sciences of the United States of America 112:E1257–E1262.

Torstrom, S. M., K. L. Pangle, and B. J. Swanson. 2014. Shedding subspecies: the influence of genetics on reptile subspecies taxonomy. Molecular Phylogenetics and Evolution 76:134–143.

Weisrock, D. W., P. M. Hime, S. O. Nunziata, K. S. Jones, M. O. Murphy, S. Hotaling, and J. D. Kratovil. 2018. Surmounting the large-genome "problem'' for genomic data generation in salamanders, p. 115–142. *In*: Population Genomics: Wildlife. P. A. Hohenlohe and O. P. Rajora (eds.). Springer International Publishing, Cham.

Wiens, J. J. 2007. The Amphibian Tree of Life. Bulletin of the American Museum of Natural History, Number 297. By D. R. Frost et al. Quarterly Review of Biology 82:55–56.

Yuan, Z.-Y., W.-W. Zhou, X. Chen, N. A. Poyarkov, Jr., H.-M. Chen, N.-H. Jang-Liaw, W.-H. Chou, N. J. Matzke, K. Iizuka, M.-S. Min, S. L. Kuzmin, Y.-P. Zhang, D. C. Cannatella, D. M. Hillis, and J. Che. 2016. Spatiotemporal diversification of the true frogs (genus *Rana*): a historical framework for a widely studied group of model organisms. Systematic Biology 65:824–842.

Zeisset, I., and T. J. C. Beebee. 2008. Amphibian phylogeography: a model for understanding historical aspects of species distributions. Heredity 101:109–119.