Secondary contact, gene flow and clinal variation between two mtDNA lineages of the Northeastern ringneck snake *Diadophis punctatus edwardsii* (Colubroidea: Dipsadidae)

FRANK M. FONTANELLA[1*], PEDRO F. VICTORIANO[2] and JACK W. SITES Jr[3]

[1]*Department of Biology, University of West Georgia, 1601 Maple Street, Carrollton, GA 30118, USA*
[2]*Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Casilla 160-C, Concepcion, Chile*
[3]*Department of Biology and Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA*

There is an emerging consensus that the intent of most species delimitation methods (SDM) is to identify evolutionarily distinct lineages. However, the criteria employed differ among SDMs depending on the perceived importance of various attributes of evolving populations. While recent work has shown that species can be delimited when the amount of gene flow between lineages is low, the accuracy of these methods declines with increasing gene flow. In this study, we examine species delimitation in the northern ringneck snake *Diadophis punctatus edwardsii* using a combined approach from morphological, molecular and ecological niche modelling. Overall, our expanded sampling inferred two well-supported mtDNA phylogroups that diverged during the Pleistocene and have undergone recent population expansions, resulting in zones of secondary contact. Ecological niche modelling of the two groups did not identify separate niche space for the groups, but instead predicted large areas of shared niche space suitable to either group. The microsatellite loci, while highly variable, failed to infer a pattern consistent with the mtDNA data and showed considerable gene flow across areas of overlapping niche space. The multivariate analyses of 18 morphometric variables failed to distinguish discrete groups consistent with the mtDNA-defined phylogroups, nor did we infer a significant correlation (positive or negative) between body size and latitude (Bergmann's rule). However, we did detect a significant pattern of decreasing head size relative to snout–vent length with latitude. Therefore, given the lack of any corroborating evidence suggesting independent evolutionary trajectories for the mtDNA phylogroups, we do not recommend the phylogroups be recognized as separate species until sampling of genetic variation is complete or a stable biological process can be identified to explain the observed genetic variation.

ADDITIONAL KEYWORDS: gene flow – mtDNA – microsatellite loci – morphology – species distribution modelling.

INTRODUCTION

Species diversification and current population structure reflect both historic and contemporary ecological and evolutionary forces often associated with past vicariant events (Hewitt, 2000). Repeated glaciations during the Pleistocene Epoch were the major force forming species pools for current local communities in North America. During times of drastic climate change such as interglacial periods, organisms often experience extinction over large portions of their ranges, survive in glacial refugia, dispersal to new locations or expand their niche space (Hewitt, 2000; Jezkova *et al.*, 2016). Many hypotheses suggest that organisms survived the Pleistocene in multiple isolated refugia, where adaptation to different environmental conditions fostered divergence and intraspecific variation (Hewitt, 2000). These processes left their signatures

*Corresponding author. E-mail: ffontane@uwg.edu

in the genomes of species, and thus, current population genetic structure can be used to infer evolutionary and demographic events within species (Avise, 2000). Following the retreat of the glaciers, populations typically depict various levels of expansion associated with recolonization of previously unsuitable habitat, often resulting in area of secondary contact or parapatry.

Secondary contact zones between closely related taxa or divergent populations allow for insights into early stages of allopatric speciation. Studies on the genetic structure of contact zones are also particularly important for understanding the complex nature of hybridization processes (Coyne & Orr, 2004; Leaché, 2011). In the absence of reproductive isolation, results of secondary contact can vary from zones of intergradation to complete admixture of parental taxa, while prezygotic (behaviour, ecological adaptations) and postzygotic (genetic incompatibilities) isolation mechanisms are expected to maintain the distinctiveness of parental gene pools via limited introgression across hybrid zones (Barton & Hewitt, 1985).

Traditionally, species delimitation has been based on morphological differences, but the inclusion of molecular and ecological data in recent years has provided more evidence for delimiting species (Sites & Marshall, 2003, 2004; Rissler & Apodaca, 2007; Leaché *et al.*, 2009). Mitochondrial DNA has been the primary and often sole source of genetic information for analyses of species level relationships and population structure, the advantages and limitations of which are well known (e.g. Zang & Hewitt, 2003; Ballard & Whitlock, 2004; Zink & Barrowclough, 2008). It is now becoming common for studies of molecular phylogeography to incorporate a diverse suite of genetic markers from both the mitochondrial and nuclear genome. For most studies that employ a diverse array of genetic markers, the populations that harbour deep splits in the mitochondrial locus often have corresponding differences in the nuclear genome (Zink & Barrowclough, 2008), but a concordant pattern of divergence in both genomes is not always observed (Funk & Omland, 2003; Chan & Levin, 2005). Therefore, testing patterns based on mtDNA with other sources of data is crucial for making robust inferences about population structure, demographic history and species limits (Brito & Edwards, 2009).

In addition to molecular and morphological data, species distribution modelling has become a useful tool for identifying and diagnosing species (Raxworthy, Ingram & Pearson, 2007; Rissler & Apodaca, 2007; Stockman & Bond, 2007; Bond & Stockman, 2008). Niche modelling is able to provide evidence for geographic isolation among populations (based on either conserved or divergent ecological niches) and considers populations to be separately evolving lineages when gene flow is deemed unlikely for the intervening geographic

regions (Wiens & Graham, 2005). Historically, the two fundamental goals of systematic biology have been the discovery of new species and the reconstruction of organismal relationships. De Queiroz (1998, 2007) has argued that there is a general agreement on the fundamental nature of species: species are separately evolving metapopulation lineages, and what differs between concepts are the criteria used to identify those lineages. While numerous criteria for the identification of lineages have been identified (reviewed in Sites & Marshall, 2003; Fujita *et al.*, 2012), it is reasonable to propose that the greater number of species criteria satisfied, the more likely it is that a group represents an identifiable biological entity on an independent evolutionary trajectory (De Queiroz, 2007).

The Northern ringneck snake *Diadophis punctatus edwardsii* is a small, abundant secretive snake native to the northeastern USA and southern Canada. The subspecies is viewed as a habitat and diet generalist, living in a diversity of habitats and feeding on a large variety of prey. Previous range-wide molecular work identified an expansive and diverse mtDNA clade ranging from southern Tennessee north through the Appalachians into southern Canada and west to central Illinois, which corresponds to the current range for the designated subspecies (Fontanella *et al.*, 2008; Fontanella & Siddall, 2009a). Although the sampling was limited, the authors found some evidence of multiple subclades with overlapping distributions, possibly a result of isolation followed by secondary contact. Although mitochondrially encoded loci have been used extensively to study snake phylogeography and population genetics (Fontanella *et al.*, 2008), their matrilineal inheritance and rapid coalescence (Avise, 2000) may obscure patterns of introgression. The combination of mtDNA with morphological characters and biparentally inherited nuclear loci provides a more powerful method of testing population genetic hypotheses and comparing matrilineal vs. nuclear genome patterns of genetic variation.

In this study, we expand upon previous mitochondrial sampling and combine morphological characters, ecological niche modelling and microsatellite loci to investigate patterns of historical demography and gene flow across a secondary contact zone in the range of *D. p. edwardsii*. We examine whether the phylogroups merit taxonomic recognition as species in light of multiple lines of evidence under the general lineage concept.

METHODS

MITOCHONDRIAL LOCUS

Genomic DNA was obtained from liver, muscle or shed skins using the Qiagen DNA extraction kit. We amplified the mitochondrial gene cytochrome *b* (1117 bp)

via PCR using previously published primers and protocols (Fontanella *et al.*, 2008). The cytochrome *b* gene region has been used successfully to examine intraspecific variation in snakes (Feldman & Spicer, 2002; Nagy *et al.*, 2004; Fontanella *et al.*, 2008). Amplifications were purified using either 2 μL of ExoSap-it (USB Corp.) per 10 μL of PCR product or AMPURE (Agentcourt) following the manufacturer's instructions. Purified PCR products were sequenced in both directions using BigDye v.3.1 Cycle Sequencing Ready Reaction (Applied Biosystems, Perkin-Elmer, CA, USA). For the cytochrome *b* gene, sequencing reactions were performed using internal sequencing primers designed specifically for *D. punctatus* (Dp-F CCTTCTGAGCAGCAACAGTAA) and (Dp-R GAAGAATCGTGTGAGGGTTGG). Reactions were purified with the CleanSeq Dye-terminator removal kit (Agentcourt) and analysed on an ABI Prism 3730 sequencer (Applied Biosystems, Perkin-Elmer, CA, USA). Sequences were assembled, edited and aligned using SEQUENCHER 4.7 (Gene Codes, Corp.), and an open reading frame was verified.

PHYLOGENETIC ANALYSES

We used jModelTest2 (Darriba *et al.*, 2012) to select the best-fit model of nucleotide change based on the Akaike information criteria (Akaike, 1973). The GTR + Γ + I (general time-reversible model with Γ-distributed among-site rate variation and with a proportion of invariant sites) model was selected as the most appropriate model of nucleotide substitution. We conducted maximum likelihood and Bayesian analyses using RAxML-VI-HPC v. 7.0.4 (Stamatakis, 2006) and MrBayes v. 3.2.6 (Ronquist *et al.*, 2012), respectively. Our RAxML analyses implemented the GTRGAMMA nucleotide substitution model with an among-site rate heterogeneity parameter partitioned across each codon position with a proportion of invariant sites. This model is equivalent to the general time-reversible model with an among-site rate heterogeneity parameter determined by ModelTest. To assess support values for the inferred relationships, we implemented 10 000 non-parametric bootstrap replicates (Stamatakis, Hoover & Rougemont, 2008).

We also conducted a Bayesian analyses using MrBAYES v.3.1.2b in which the GTR + Γ + I model was applied to the gene region. Each Markov chain was started from a random tree and run for 1.0×10^7 generations with every 1000th tree sampled from the chain. Stationarity was checked as suggested in Nylander *et al.* (2004). All sample points prior to reaching the plateau phase were discarded as 'burn-in' and the remaining trees combined to find the *a posteriori* probability estimate of phylogeny. Branch lengths were estimated as means of the posterior probability density. Trees were visualized using the FigTree v1.1.2 program, available at http://tree.bio.ed.ac.uk/software/figtree/. For the outgroup taxa, we chose individuals from both distantly and closely related lineages (Piedmont and Western Kentucky) as defined by Fontanella & Siddall (2009a).

To infer the date of origin for each lineage without relying on a molecular clock and considering uncertainty in tree topology and branch length (i.e. 'the relaxed phylogenetics' method), we used BEAST v1.8.3 (Drummond *et al.*, 2012). Phylogenetic estimates were constructed under the GTR + Γ + I model with an uncorrelated lognormal tree prior with a constant population size prior. A mean calibration point of 0.376 Mya was placed at the root of ingroup with a lognormal SD of 0.3 producing a 95% credible sampling interval from 0.21 to 0.564 Myr (Fontanella & Siddall, 2009a). Analyses were run for 20 million generations and sampled every 1000th iteration following a pre-burn-in of 2000 generations.

Haplotype diversity (Hd), nucleotide diversity (π) and the average number of pairwise differences (*k*) were calculated to quantify DNA polymorphisms in each lineage using DNAsp v3.0 (Roza & Roza, 1999). Mismatch distributions were conducted to further investigate the possibility of historic demographic change. The fit of the observed data (R_2) (Ramos-Onsins & Rozas, 2002) was compared using the sum of squares deviations between observed and expected data estimated from 10 000 coalescent simulations in DnaSP.

MICROSATELLITE LOCI

We selected eight microsatellite primers specifically designed for *D. punctatus* from Fontanella & Siddall (2009b). Each forward primer was augmented on the 5′ end with an M13 forward sequence (CACGACGTTGTAAACGAC). This modified primer was then used in combination with the 5′ FAM fluorescently labelled forward M13 primer in subsequent amplification reactions. Thus, final PCR reactions consisted of 1 μL of DNA, 0.25 μL of unmodified reverse primer (10×), 0.025 μL of M13 tailed forward primer (10×), 0.45 μL of 6-FAM (Geneworks) fluorescently labelled M13 primer (10×), 1.25 μL of PCR Buffer, 1.25 μL of 25 mM $MgCl_2$, 0.2 μL of 10 mM dNTP mixture, 0.1 μL of Taq and 8.0 μL of water for a total volume of 12.5 μL. The thermal profile was 95 °C for 2 min followed by 30 cycles of 94 °C for 30 s, the primer-specific annealing temperature and time and 68 °C for 30 s followed by a 4-min extension.

Depending on the overall strength of the amplification reaction, PCR were diluted 1:10 or 1:30 with water into a single plate for genotyping. One microlitre of the bulk PCR dilution was added to a plate containing

0.09 µL of the GeneScan 500 LIZ genotyping standard and 9.91 µL of HiDi Formamide (Applied Biosystems). Reactions were genotyped on an ABI 3730XL automated sequencer and automatically scored using GENOTYPER 3.7 (Applied Biosystems). Scoring of microsatellite alleles was verified by eye for each sample. MICRO-CHECKER software (Oosterhout *et al.*, 2004) was used to assess the presence of genotyping errors, such as non-amplified alleles, short allele dominance and scoring of stutter peaks.

MICROSATELLITE ANALYSIS

For each microsatellite locus, we calculated the total number of alleles in FSTAT v.2.9.3.2 (Goudet, 1995) and used Arlequin 3.11 (Excoffier, Laval & Schneider, 2005) to detect significant departures from Hardy–Weinberg expectations for all pairs of loci, mean heterozygosity (Theta H; Ohta & Kimura, 1973), and to test for linkage disequilibrium (non-random association between loci) among all pairs of loci within each lineage. We used default parameters in FSTAT and Arlequin for all Markov-chain tests and permutations. Analysis with the MICRO-CHECKER (Oosterhout *et al.*, 2004) software showed no evidence that null alleles, stuttering or large allele dropout affected any of the loci, and rescoring of samples confirmed that alleles were being sized consistently among sequencing runs; consequently, all loci were included in all analyses.

To assess patterns of genetic structure and to test for levels of shared ancestry between lineages from microsatellite data, we used Bayesian clustering software: STRUCTURE 2.3.4 (Pritchard, Stephens & Donnelly, 2000) and GENELAND 4.0.3. To assign individuals to genetic groups (K), ten replicate runs were conducted in STRUCTURE for each value of K, which varied from 1 to 9. Because we were interested in examining potential gene flow between lineages, we used the admixture model with correlated allele frequencies for each run with no prior placed on population origin. These parameters allow for the determination of hybrids by estimating the fraction of ancestry for each individual from each cluster. Each run included a burn-in of 500 000 followed by 2 000 000 Markov chain Monte Carlo (MCMC) iterations. The most likely number of clusters was estimated using the ΔK statistic of Evanno, Regnaut & Goudet (2005). A final run including a burn-in of 1 million followed by 5 million MCMC iterations was run using the K value determined by the ΔK statistic. We used GENELAND to incorporate fine-scale geographical information into the analysis of population structure. Five runs were performed for the spatially explicit model, and for each run, the posterior probability of subpopulation membership was computed for each pixel of the spatial domain (100 × 100 pixels). The MCMC was run for 10 million generations, thinning was set at 100 with the initial 20% discarded as burn-in. We implemented the spatial model and correlated allele frequencies using $K = 2$ determined by the ΔK statistic.

ECOLOGICAL NICHE MODELLING

The realized environmental niche of a species can be estimated from presence-only data with high precision by extracting niche dimensions from spatial information on the distribution of environmental parameters (Nix, 1986). We used the maximum entropy model implemented in the program MAXENT v.3.3.1 (Phillips, Anderson & Schapire, 2006) to predict where the lineages of *D. p. ewdwardsii* are most likely to occur under current climatic conditions. MAXENT generates ecological niche models using presence-only records, contrasting them with pseudo-absence data sampled from the remainder of the study area. We chose this approach because of its overall better performance with presence-only data and under conditions of small sample size (Elith *et al.*, 2006). Museum samples were assigned to a particular lineage if they placed inside the known range of a mtDNA haploclade distribution. Samples that fell outside the currently known ranges were omitted. We recognize that the omission of these samples may bias the range estimation for lineages.

For the contemporary niche predictions, we used the 19 bioclimatic variables from the WorldClim data set (v. 1.4) with a resolution of 30-s resolution (Hijmans *et al.*, 2005) for the combined morphological and molecular samples ($N = 463$). Bioclimatic variables were derived from monthly temperature and precipitation layers and represent biologically meaningful properties of climate variation (Hijmans *et al.*, 2005; Waltari *et al.*, 2007).

The models were run with the default convergence threshold (10^{-5}), with a 1000 iterations and 25% of localities for model training. The program selected both suitable regularization values and functions of environmental values automatically, based on considerations of sample size. Because the samples removed for model training can affect the overall predicted distribution, we generated ten models for each lineage and averaged the results using the cross-validation option. MAXENT outputs a continuous probability value (logistic values), which is an indicator of relative suitability for the species. Model clamping was checked with the 'fade by clamping' option available in Maxent v 3.3.1. To determine the threshold value for each projection, we used the minimum training value averaged over the ten runs.

To estimate niche overlap, we used the statistic D, developed by Schoener (1968), and applied it to environmental niche models by Warren, Glor & Turelli (2008). D describes the difference between two niche

models in the predicted probability of presence across a study area, scaled from 0 (no overlap) to 1 (identical models). The statistic is useful because it does not require the conversion of continuously distributed probability of presence values into a binary prediction of cell occupancy via the arbitrary selection of a threshold value defining taxon presence. Measures of niche overlap can be sensitive to choice of threshold (Warren *et al.*, 2008). We calculated a mean estimate of niche overlap (D_m) as the average D across ten replicate runs.

MORPHOLOGICAL DATA

To investigate whether independent, morphologically distinct lineages exist within *D. p. edwardsii*, we examined 18 external characters collected from 321 samples [176 from the Northern clade; 145 from the Southern clade (list of characters in Supporting Information, Text S1)]. All measurements were made using PITTSBURGH digital calipers (± 0.002 mm). In an exploratory step, we obtained correlations among each variable and total body size. Given a significant association between body size and all the other variables, and in order to test statistical differences in morphology among lineages (Northern and Southern clades), we carried out two way covariance analysis (Ancovas), considering the factors lineage and sex, and using total body size as a covariable. After the Ancovas, we selected those variables with significant differences among lineages. With the selected variables after Ancovas, and in order both to reduce the number of variables and to detect the occurrence of groups consistent with each of two molecular clades, we carried out a principal component analysis (PCA). For PCA we used the residuals obtained from correlations among each selected variable and total length. After the PCA analysis, we reduced the variable list by selecting those with higher loads on the two first principal components, and we carried out a discriminant analysis (DA) and multivariate numerical functions based on residuals of correlations obtained as for the PCA. Based on the numerical functions, we estimated the success of the functions in predicting the right assignment for each individual to each clade and sex. Given the absence of discrete morphological groups for each clade, and the low percentage of success assignments of discriminant functions (see Supporting Information, Table S2), we estimated the probable geographical (latitudinal) pattern for morphology, and its association with the lineages. For this analysis, we obtained two correlations using latitude as an independent variable. For the first correlation, the canonical value for each individual was used as the dependent variable, and a second correlation was used for the posterior probability of assignment of each individual to each lineage as a dependent variable. Individual canonical values and posterior probabilities summarize in one value, the morphology of each individual, and it can be used to test patterns of association of morphology with geography. All the multivariate analyses were conducted by using STATISTICA v7.0 (StatSoft, 2004).

RESULTS

PHYLOGENETIC ANALYSIS

We obtained 142 specimens that represented the majority of the known geographic range for *D. p. edwardsii* (Fig. 1). All sequences were submitted to GenBank (Accession Numbers KY508698–KY508797). The translated amino acid sequences did not contain stop codons or indels. BLAST results indicated a high sequence similarity to other lineages of the *Diadophis* genus as well as other Dipsadinae groups. The presence of an open reading frame suggests that the amplicon was a fragment from the functional mitochondrial cytochrome *b* gene and not a nuclear pseudogene or DNA contaminant.

Because both the ML and BI analysis produced highly congruent estimates of the phylogeographic patterns, only the Bayesian consensus phylogram is presented with the posterior probabilities and nonparametric bootstrap values for the shared branches (Fig. 2). The Bayesian analysis produced a 50% majority-consensus tree with a harmonic mean of −4166.6 following a burn-in of 10 000 generations. The analyses inferred two well-supported clades (Northern and Southern, respectively) with poorly supported internal nodes and short branch lengths (Fig. 2). The Northern lineage has a broad range extending across most of the northeastern USA east of the Mississippi River and excluded only from the Southern Blue Ridge, Cumberland and Southern Ridge Valley ecoregions (Fig. 1). The Southern lineage is mostly restricted to west of the mid-Atlantic coastal plain and extends to central Illinois and north into the Appalachian Mountains in central West Virginia (Fig. 1).

Results of the dating analysis suggest that the time to the most recent common ancestor (TMRCA) for the two clades date to the Pleistocene with a mean estimate of 364 thousand years ago, while the TMRCA for each lineage occurred during the early Pleistocene (Fig. 2). The effective sample size was greater than 200, suggesting that the 20 million generations were sufficient to determine age range for each lineage.

Haplotype diversity (Hd) and the mean number of pairwise differences (k) were high for the Northern and Southern lineages, while the nucleotide diversity (π) was low within each lineage. Additionally, across all samples, π and k were approximately double those of the two main clades (Table 1). The mismatch
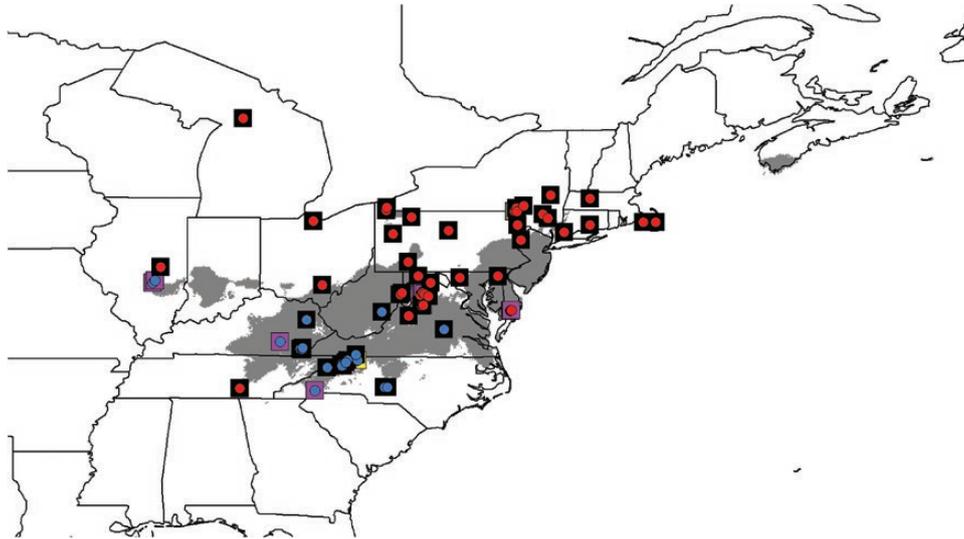
Figure 1. Map of the eastern USA showing location of samples used in this study. Circles depict individuals assigned to clades from mitochondrial DNA (red = Northern; blue = Southern). Squares depict individuals assigned to populations from microsatellite loci (black = Northern; Southern = purple). Yellow squares depict individuals of mixed ancestry. Grey area represents regions of overlapping predicted niche space.

distributions of pairwise nucleotide differences were unimodal and Tajima's D and Fu's F_s were significantly negative ($P = 0.05$), all of which suggests a recent population expansion for each lineage (Table 1; Fig. 3).

MICROSATELLITE LOCI

Genotypes for 8 microsatellite loci were resolved for 142 snakes, and the number of alleles per locus ranged from 5 to 30, with a greater number of alleles present in the Northern lineage. Genetic diversity was high across all loci in both lineages and was significantly different between groups ($P > 0.05$). Theta ranged from 6.52 to 43.6 in the Southern lineage and from 9.77 to 150.3 in the Northern lineage (Table 2). We found no evidence of linkage disequilibrium or departures from Hardy–Weinberg ratios, and none of the loci were found to be under selection.

The ΔK statistic indicated that $K = 2$ is the best supported number of genetic clusters to fit the data. Neither clustering method recovered a pattern consistent with the mtDNA tree, instead inferring high levels of gene flow between lineages (Fig. 4). Both analyses inferred one broadly distributed cluster with high posterior probability that included nearly all sampling sites, while the other cluster consisted of three sampling sites (Fig. 4, Supporting Information, Figs S3 and S4).

ECOLOGICAL NICHE MODELLING

The present bioclimatic niche ranges for the lineages of *D. p. edwardsii* are shown in Figure 5. Model validation was conducted by calculating the area under the curve

(AUC), which reflects the model's ability to distinguish between presence records and random background points (Hanley & McNeil, 1982). AUC values range from 0.5 for models with no predictive ability, to 1.0 for models with perfect predictive ability. According to Swets (1988), AUC values > 0.9 are considered to have 'very good', > 0.8 'good' and > 0.7 'useful' discrimination abilities. The AUC scores were relatively high for both the Northern (0.91 ± 0.04) and Southern (0.84 ± 0.09) lineages. The predicted region of suitable habitat for Northern clade closely matched the distribution of the samples along their latitudinal gradient; however, there was a large area of suitable habitat predicted west of the mtDNA clade sampling area. This region of suitable habitat extends west of the Mississippi River, which has been shown to be a strong barrier to dispersal and gene flow for terrestrial organisms (Soltis *et al.*, 2006). The predicted region of suitable habitat for the Southern clade depicted a broader range extending north along the eastern coastline into southeastern Pennsylvania and New Jersey (Fig. 5). Despite the relatively good fit of each model, there was a large area of shared suitable habitat in the central portion of the eastern USA (Fig. 5). Schoener's (1968) metric of niche overlap (D) inferred substantial niche overlap between the two phylogroups ($D = 0.352 \pm 0.014$), indicating that the predicted niches are not climatically distinct.

MORPHOLOGICAL DATA

After the Ancovas, we selected those variables with significant differences between clades. A total of 13 morphometric variables were selected for the PCA.
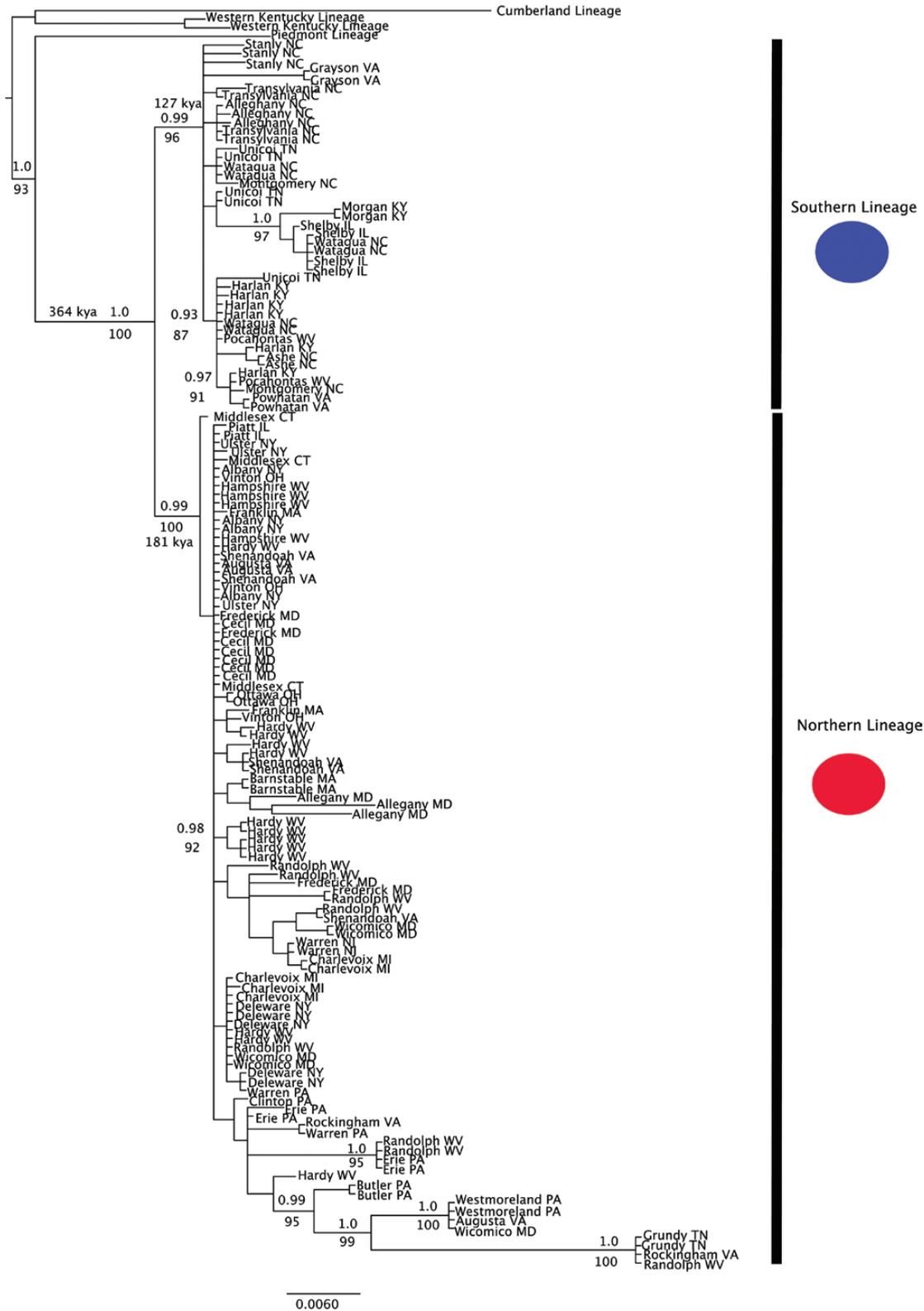
Figure 2. Bayesian 50% majority-rule consensus tree for the 142 *Diadophis punctatus edwardsii* samples and outgroup taxa. Support values (Posterior probabilities above branches, non-sparametric bootstrap proportions below) are shown for nodes with a posterior probability > 95%. Time to most recent common ancestor given in thousands of years (kya). Circles colours correspond to those in Figure 1.

Table 1. Sample size (N), haplotype diversity (Hd), nucleotide diversity (pi), average number of pairwise differences (K) and results of Tajima's *D* and Fu's F_s statistic for each lineage calculated for all mtDNA sites

Lineage	N	Hd	pi	k	Tajima's D	Fu's F_s	r_g	R_2
Northern	99	0.952	0.007	8.69	−2.527*	−23.17*	0.007	0.042
Southern	43	0.977	0.007	7.97	−2.045*	−11.87*	0.015	0.049
Total	142	0.976	0.013	14.16	−2.33*	−35.35*	0.004	0.036

The raggedness statistic (r_g) and the Ramos-Orsins and Rozas R_2 statistic are reported for the mismatch distributions.
*Tests that were significantly different at a *P* = 0.05.
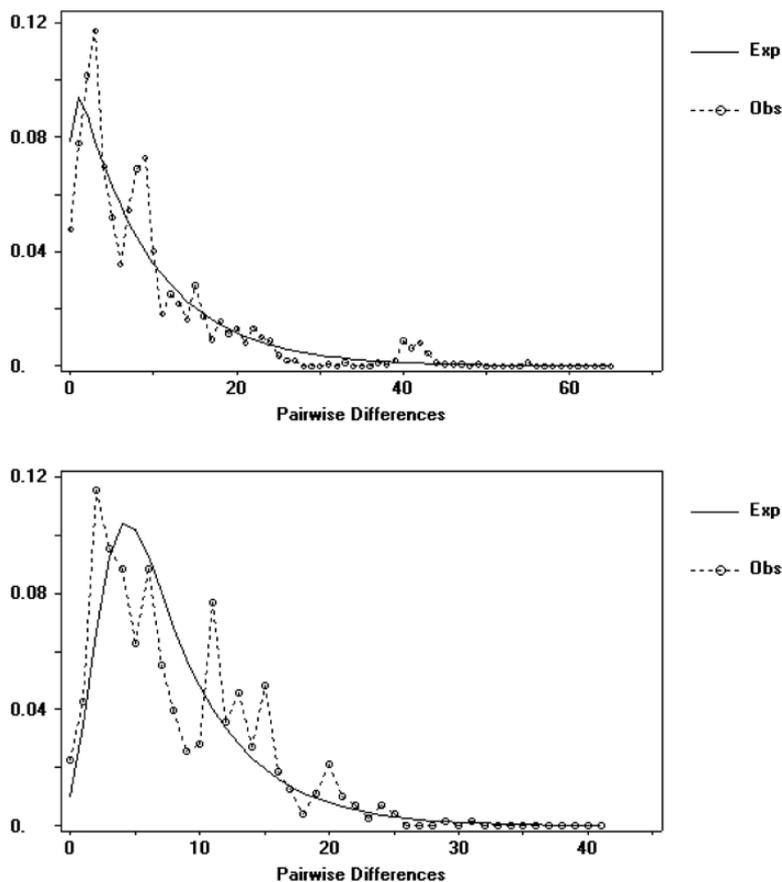
Figure 3. Mismatch distributions depicting the demographic history for the Northern lineage (A) and the Southern lineage (B). Open circles represent the observed distribution of pairwise differences and the solid line represents the expected distribution assuming population expansion.

Given the significant differences for sex for all of these variables, although with no interactions with the factor lineage, we considered sex for the future analyses too. In the PCA, Factor 1 explained 49.6% of the total variation and the sum of Factors 1 and 2 explained 57.7%. Although the total variance explained by this analysis was high, the distribution of individual values in the two factors did not conform to discrete groups consistent with the mtDNA phylogroups. However, a trend to separate each genealogical group was concordant with the first component (Supporting Information, Fig. S5). Based on the highest loads, we selected ten variables (Supporting Information, Table S6) for the DA. From the DA plot no completely discrete groups were obtained (Fig. 6), but, independent from sex, a clear trend associated with each lineage was observed. Canonical functions were obtained not in order to predict discrete groups, but to analyse the geographical pattern of morphology. A high significant correlation was obtained both for the association between canonical values and latitude, and posterior probability of assignment to each

Table 2. Summary statistics of the eight microsatellite loci for the two lineages of *Diadophis punctatus edwardsii*

Locus	Number of alleles	Southern	Lineage	$N = 43$	Number of alleles	Northern	Lineage	$N = 99$
		H_o	H_e	H		H_o	H_e	H
E4	8	0.79	0.86	27.7	8	0.87	0.85	23.9
H2	7	0.79	0.74	7.05	9	0.78	0.77	9.77
17	16	0.90	0.87	31.4	30	0.90	0.94	150.3
26	12	0.81	0.85	21.7	15	0.86	0.89	47.2
25	11	0.90	0.89	43.6	12	0.88	0.85	23.3
19	7	0.81	0.73	6.52	11	0.77	0.82	15.4
H1	5	0.76	0.70	6.87	17	0.88	0.91	73.2
E5	7	0.76	0.75	11.0	9	0.78	0.82	18.4

Number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; H, theta where $\theta = 4 N_e \mu$ for diploids (Ohta & Kumura, 1973).
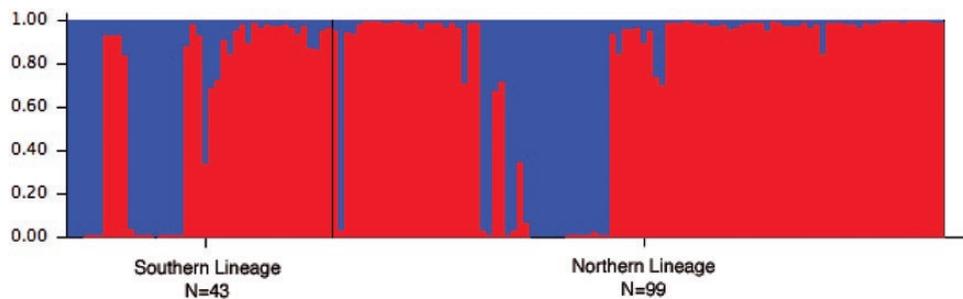
Figure 4. Bar plot output from STRUCTURE. The length of coloured bars represents the fractional assignment of individuals to each of $K = 2$ genetic clusters: red = Northern mtDNA lineage, blue = Southern mtDNA lineage.

lineage and latitude (Fig. 7). In general, the values of morphometric variables (mainly head characters), and relative to snout–vent length (SVL), showed a negative trend with latitude. After DA, individuals at the northernmost latitudes, were assigned with a higher posterior probability to the Northern molecular lineage with increasing latitude, while the opposite trend was observed for the Southern lineage.

DISCUSSION

Our genetic, morphological and species distribution modelling results provide strong evidence that the two well-supported mtDNA phylogroups in this species do not conform to or identify as well-defined evolutionary lineages that have undergone adaptive differentiation. Our phylogenetic analyses of mtDNA haplotypes recovered two well-supported clades (Northern and Southern, respectively) that diverged during the Pleistocene and presumably expanded with negligible gene flow when coming to occupy their current geographical range and establishing secondary contact (Lougheed *et al.*, 2013) (Fig. 2). Therefore, the signature reflected in the mtDNA is likely due to populations that experienced a brief period of geographic

isolation followed by population expansion and secondary contact, leading to genetic introgression and dissociated patterns between the nuclear and mitochondrial genomes.

The Northern lineage has a broad range extending across most of the northeastern USA east of the Mississippi River and is excluded only from the Southern Blue Ridge and Cumberland and Southern Ridge Valley ecoregions (Fig. 1). The Southern lineage is mostly limited at its eastern edge by the mid-Atlantic coastal plain and extends west to central Illinois and north into the Appalachian Mountains in central West Virginia (Fig. 1). While we did not find any areas where Northern and Southern haplotypes co-occurred, there was an area of broad overlap between the two clades. For instance, several Southern haplotypes occur at more northern latitudes than several Northern haplotypes and vice versa, particularly west of the Appalachian Mountains (Fig. 1). Contact zones throughout this region are not uncommon and have been observed in a variety of animals including ants (Menke *et al.*, 2010), mosquitos (Emerson *et al.*, 2010) and mammals (Hayes & Harrison, 1992).

In instances where there has been a lack of an obvious physical geographic boundary separating phylogroups, the incorporation of ecological niche
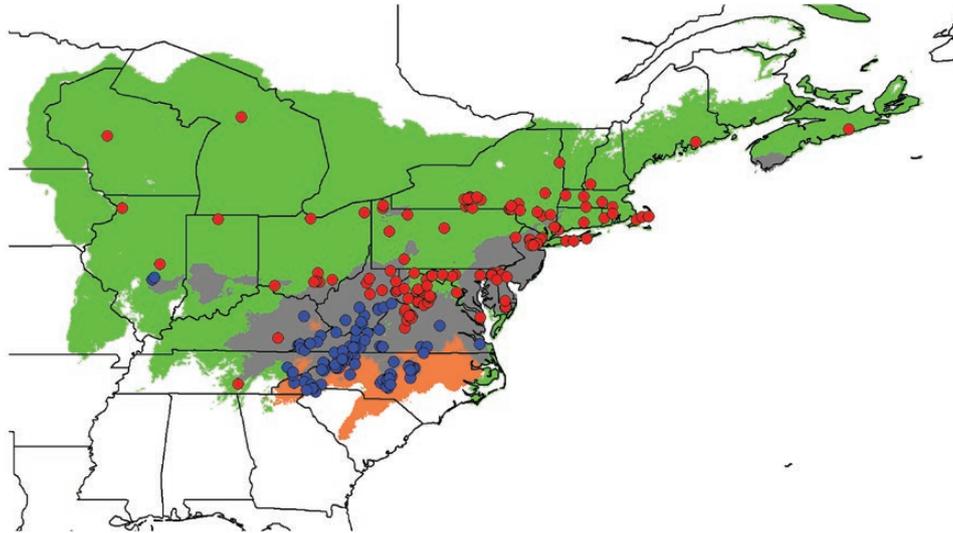
Figure 5. Ecological niche predictions for the Southern and Northern lineages based on the combined morphological and molecular samples. Symbols correspond to those in Figure 2. Grey area represents area of shared niche space.
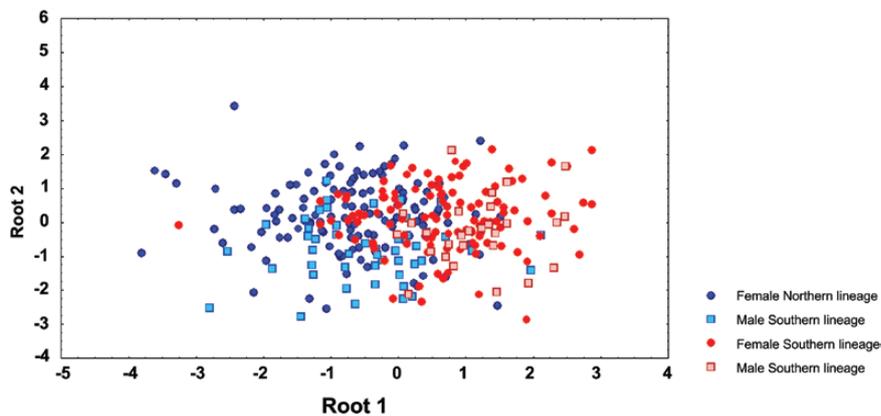
Figure 6. Discriminant function plot for two lineages of *Diadophis punctatus edwardsii* based on ten morphometric variables. Blue circles: females Northern lineage, light blue squares: males Northern lineage, red circles: females Southern lineages, pink squares: males Southern lineage.

modelling has provided evidence for geographic isolation among populations and has been used to recognize separately evolving lineages when gene flow is deemed unlikely for the intervening geographic regions (Wiens & Graham, 2005). However, our species distribution modelling failed to recover separate niche space and instead predicted large areas of overlapping niche space suitable for either phylogroup. Schoener's *D* statistic failed to infer climatically distinct predicted niches for either group and the area of predicted sympatric niche space is significantly larger ($P < 0.001$) than the predicted niche space available to the Southern phylogroup (grey areas) (Fig. 5), suggesting that the two phylogroups are not isolated by physical or ecological barriers. In a simulated study, Irwin (2002) showed that deep phylogeographic splits

can be discordant with known geographic boundaries for species that have limited dispersal capabilities and small population sizes. However, we do not suggest that the phylogeographic structure inferred from our data is a result of this phenomenon. While the life history, demography and ecology of ringneck snakes vary across their range (Fitch, 1975; Blanchard, Gilreath & Blanchard, 1979), the information available for eastern ringneck snakes suggests relatively large population sizes with moderate dispersal capabilities of over 1.6 km (Blanchard *et al.*, 1979).

Large areas of overlapping suitable niche space could provide corridors for dispersal and migration, resulting in past hybridization, intergrade zones or contemporary gene flow between populations undergoing expansion (De Le Torre, Roberts & Aitken, 2014). The
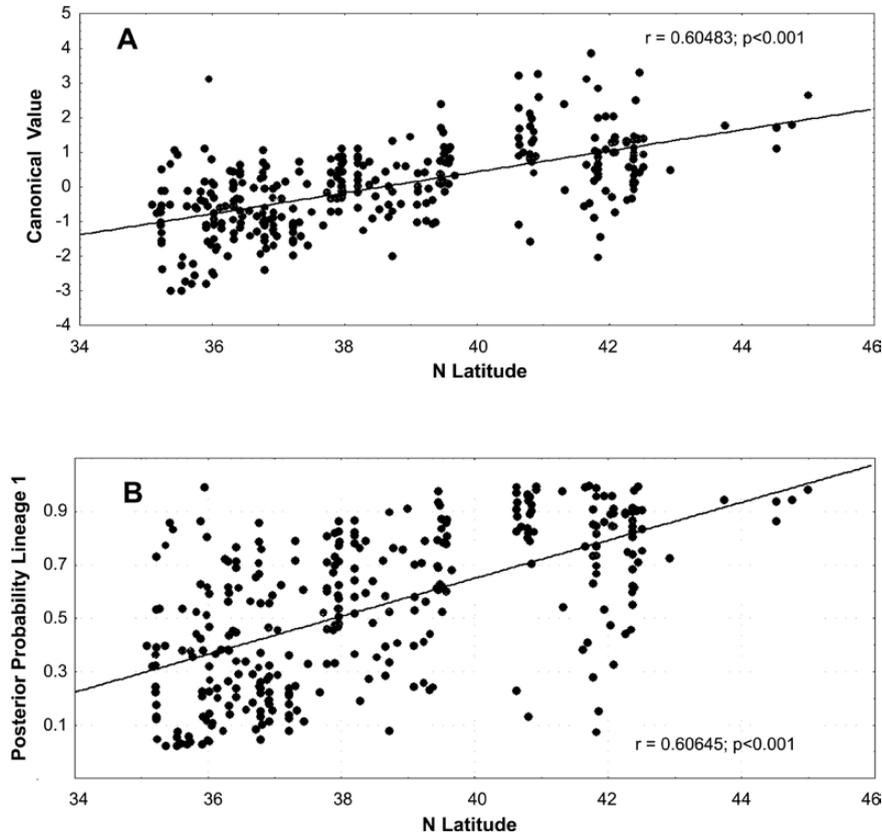
Figure 7. Correlations for (A) individual canonical values and (B) posterior probability of assignment to the lineage 1, and latitude, based on a discriminant analysis using ten morphometric variables.

population-level analyses on the mtDNA phylogroups show statistically significant results and unimodal mismatch distributions depicting recent population expansion (Fig. 3; Table 1) resulting in zones of secondary contact. Additionally, tree topologies with shallow branches, little internal resolution and low sequence divergence further suggest recent population expansion from refugial areas (Fontanella *et al.*, 2008).

The failure of the microsatellite loci to distinguish the two phylogroups provides evidence for a lack of genome wide differentiation among snakes belonging to different mtDNA lineages (Fig. 4). This pattern could be the result of a long-standing lack of differentiation among nuclear genes that predates the onset of gene flow between lineages (e.g. incomplete lineage sorting) or it could be a consequence of contemporary gene flow that occurred in this region (Gibbs *et al.*, 2006). While these results suggest there has been substantial differentiation at the peripheral region of the Northern phylogroup, the locations of the introgressed individuals depict a broad intergrade zone that is confined to overlapping predicted niche space, suggesting that the size and shape of the zone may be a result of exogenous factors (Barton & Hewitt, 1985) (Fig. 1).

The PCAs for the 18 morphometric variables failed to distinguish discrete groups consistent with the mtDNA-defined phylogroups for either sex or as a whole (Fig. 7), nor did we detect a pattern of increasing body size with latitude consistent with Bergmann's rule (Bergmann, 1847). While this well-known ecogeographic pattern has been shown to be valid for mammals and birds (Meiri & Dayan, 2003), but snakes often reverse this pattern (Ashton & Feldman, 2002; Reed, 2003; Feldman & Meiri, 2014). However, we did detect a significant trend ($P < 0.001$) in head shape with increasing latitudes for both the canonical values ($r = 0.605$), and the posterior probability assignment of each lineage ($P < 0.607$) (Fig. 7). In general, this trend was associated to smaller head characters size relative to SVL towards higher latitudes.

CONCLUSION

Older lineage-based species concepts have been used to recognize species based on reciprocal monophyly of gene genealogies (Cracraft, 1983; Donoghue, 1985; Baum & Shaw, 1995), but this method has been criticized for several problems. For instance, gene trees

may not always be congruent with species trees (Moore, 1995). Additionally, the matrilineal mode of inheritance for mtDNA results in a nested phylogenetic pattern of descent even among interbreeding lineages (Davis, 1996). Therefore, even a gene tree that accurately represents the phylogenetic history is of limited value in predicting gene flow between closely related lineages (Mallet, 2001; Schluter, 2001). Since monophyly exists at all levels in a phylogeny, the level to which species should be recognized is unclear.

Under the diagnosability criterion, species are identified as 'the smallest aggregation of populations or lineages diagnosable by a unique combination of character states in comparable individuals…' (Nixon & Wheeler, 1990, p. 211). Our multivariate analyses of the morphological characters failed to recover diagnosable groups (Fig. 6) for either sex, thus neither lineage conforms to the diagnosability criterion.

The genotypic cluster criterion is based on the idea that species are 'distinguishable groups of individuals that have few or no intermediates when in contact' (Mallet, 1995). We used the Bayesian MCMC method implemented in STRUCTURE and GENELAND to quantify the degree of admixture or the absence of intermediates between phylogroups. The results failed to recover genotypic clusters concordant with the mtDNA phylogroups and also depicted high levels of gene flow within and across areas of secondary contact (Fig. 6).

Finally, the ecological species criterion predicts that a 'species is a lineage (or closely related set of lineages) which occupies an adaptive zone minimally different from that of any other lineage in its range and which evolves separately from all lineages outside its range' (Van Valen, 1976, p. 223). Again, the ecological niche modelling failed to identify any separate or restricted areas for either lineage, and instead predicted large areas of shared niche space that encompass the vast majority of the range predicted for the Southern phylogroup (Fig. 5).

Therefore, given the lack of any corroborating evidence suggesting independent evolutionary trajectories for the mtDNA phylogroups, combined with the concerns associated with the use of lineage-based methods, particularly for when mtDNA is the sole marker, we do not recommend the phylogroups be recognized as separate species. Our study adds to the growing list of recent studies documenting contact zones between maternal lineages of animal systems where there is strong evidence of discordance between biogeographic patterns identified in mitochondrial DNA and those observed in the nuclear genome (Toews & Brelsford, 2012).

ACKNOWLEDGEMENTS

REFERENCES

Akaike H. 1973. Information theory and an extension of the maximum-likelihood principle. In: Petrov BN, Csaki F, eds. *Second International Symposium on Information Theory, Tsahkadzor, Armenia, USSR*. Budapest: Akademiai Kiado, 267–281.

Ashton KG, Feldman CR. 2002. Bergmann's rule in nonavian reptiles: turtles follow it, lizards and snakes reverse it. *Evolution* **57:** 1151–1163.

Avise JC. 2000. *Phylogeography: the history and formation of species*. Cambridge: Harvard University Press.

Ballard JW, Whitlock MC. 2004. The incomplete natural history of mitochondria. *Molecular Ecology* **13:** 729–744.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16:** 113–148.

Baum DA, Shaw KL. 1995. Genealogical perspectives on the species molecular problem. In: Hoch PC, Stephenson AG, eds. *Experimental and molecular approaches to plant biosystematics*. St. Louis: Missouri Botanical Garden, 289–303.

Bergmann KGL. 1847. Ueber die Verhaltmisse der Warmeokonomie der thiere zu ihrer grosse. *Gottinger Studien* **3:** 595–708.

Blanchard FN, Gilreath M, Blanchard F. 1979. The eastern ring-neck snake (*Diadophis punctatus edwardsii*) in northern Michigan. *Journal of Herpetology* **13:** 377–402.

Bond JE, Stockman AK. 2008. An integrative method for delimiting cohesion species: finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology* **57:** 628–646.

Brito PH, Edwards SV. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* **135:** 439–455.

Chan KM, Levin SA. 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59:** 720–729.

Coyne JA, Orr HA. 2004. *Speciation*. Sunderland: Sinauer Associates.

Cracraft J. 1983. Species concepts and speciation analysis. *Current Ornithology* **1:** 159–187.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

Davis JI. 1996. Phylogenetics, molecular variation and species concepts. *Bioscience* **46:** 502–511.

De Le Torre AR, Roberts DR, Aitken SN. 2014. Genome wide admixture and ecological niche modelling reveal the maintenance of species boundaries despite long history of interspecific gene flow. *Molecular Ecology* **23:** 2046–2059.

De Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH, eds. *Endless forms: species and speciation*. New York: Oxford University Press, 57–75.

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

Donoghue MJ. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist* **88:** 172–181.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soveron J, Williams S, Wisz MS, Zimmerman NE. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29:** 129–151.

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM. 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 16196–16200.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611–2620.

Excoffier L, Laval G, Schneider S. 2005. Arlequin suite version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

Feldman A, Meiri S. 2014. Australian snakes do not follow Bergmann's rule. *Evolutionary Biology* **41:** 327–335.

Feldman CR, Spicer GS. 2002. Mitochondrial variation in sharp-tailed snakes (*Contia tenuis*): evidence of a cryptic species. *Journal of Herpetology* **36:** 648–655.

Fitch HS. 1975. *A demographic study of the ringneck snake (Diadophis punctatus) in Kansas*, Vol. 62. Lawrence: University of Kansas, Museum of Natural History, Miscellaneous Publication, 1–53.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution* **27:** 480–488.

Fontanella FM, Feldman CR, Siddall ME, Burbrink FT. 2008. Phylogeography of *Diadophis punctatus*: extensive lineage diversity and repeated patterns of historical demography in a trans-continental snake. *Molecular Phylogenetics and Evolution* **46:** 1049–1070.

Fontanella FM, Siddall ME. 2009a. Evaluating hypotheses on the origin and diversification of the ringneck snake *Diadophis punctatus* (Colubridae: Dipsadinae). *Zoological Journal of the Linnean Society* **158:** 629–640.

Fontanella FM, Siddall ME. 2009b. Isolation and characterization of 14 polymorphic microsatellite loci in the ringneck snake *Diadophis punctatus* (Colubridae: Dipsadinae). *Conservation Genetics* **11:** 1193–1195.

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34:** 397–423.

Gibbs HL, Corey SJ, Blouin-Demers G, Prior KA, Weatherhead PJ. 2006. Hybridization between mtDNA-defined phylogeographic lineages of black ratsnakes (*Pantherophis* sp.). *Molecular Ecology* **15:** 3755–3767.

Goudet J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86:** 485–486.

Hanley JA, McNeil BJ. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143:** 29–36.

Hayes JP, Harrison RG. 1992. Variation in mitochondrial DNA and biogeographic history of woodrats (*Neotoma*) of the eastern United States. *Systematic Biology* **41:** 331–344.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907–913.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25:** 1965–1978.

Irwin DE. 2002. Phylogeographic breaks without geographic barriers to gene flow. *Evolution* **56:** 2383–2394.

Jezkova T, Jaeger JR, Oláh-Hemmings V, Jones KB, Lara-Resendiz RA, Mulcahy DG, Riddle BR. 2016. Range and niche shifts in response to past climate change in the desert horned lizard (*Phrynosoma platyrhinos*). *Ecography* **39:** 437–448.

Leaché AD. 2011. Multi-Locus Estimates of Population Structure and Migration in a Fence Lizard Hybrid Zone. *PLoS ONE* **6:** e25827.

Leaché AD, Koo MS, Spencer CL, Papenfuss TJ, Fisher RN, McGuire J. 2009. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). *Proceedings of the National Academy of Sciences of the United States of America* **106:** 12418–12423.

Lougheed SC, Campagna L, Davila JA, Tubaro PL, Lijtmaer DA, Handford P. 2013. Continental phylogeography of an ecologically and morphologically diverse Neotropical songbird, *Zonotrichia capensis*. *BMC Evolutionary Biology* **13:** 1–16.

Mallet J. 1995. A species definition for the modern synthesis. *Trends in Ecology & Evolution* **10:** 294–299.

Mallet J. 2001. Species, concetps of. In: Levin SA, ed. *Encyclopedia of biodiversity*, *Vol. 5*. New York: Academic Press, 427–440.

Meiri S, Dayan T. 2003. On the validity of Bergmann's rule. *Journal of Biogeography* **30:** 331–351.

Menke SB, Booth W, Dunn RR, Edward CS, Vargo EL, Silverman J. 2010. Is it easy to be urban? Convergent success in urban habitats among lineages of a widespread native ant. *PLoS ONE* **5:** e9194.

Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear gene trees. *Evolution* **49:** 718–726.

Nagy ZT, Lawson R, Joger U, Wink M. 2004. Molecular systematics of racers, whipsnakes and relatives (Reptilia: Colubridae) using mitochondrial and nuclear markers. *Journal of Zoological Systematics and Evolutionary Research* **42:** 223–233.

Nix HA. 1986. A biogeographic analysis of Australian elapid snakes IN Longmore R. ed Atlas of Elapid snakes in Australia. *Australian Flora and Fauna* **8:** 4–15.

Nixon KC, Wheeler QD. 1990. An amplification of the phylogenetic species concept. *Cladistics* **6:** 211–223.

Nylander JA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey J. 2004. Bayesian phylogenetic analysis of combined data. *Systematic Biology* **53:** 47–67.

Ohta T, Kimura M. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research* **22:** 201–204.

Oosterhout CV, Hutchinson WF, Wills DP, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources* **4:** 535–538.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modeling* **190:** 231–259.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Ramos-Onsins SE, Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19:** 2092–2100.

Raxworthy CJ, Ingram CM, Pearson RG. 2007. Species delimitation applications for ecological niche modeling: a review and empirical evaluation using *Phelsuma* day gecko groups from Madagascar. *Systematic Biology* **56:** 907–923.

Reed RN. 2003. Interspecific patterns of species richness, geographic range size, and body size among New World venomous snakes. *Ecography* **26:** 107–117.

Rissler LJ, Apodaca JJ. 2007. Ecological niche models and phylogeography help uncover cryptic biodiversity. *Systematic Biology* **56:** 924–942.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA,

Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Roza J, Roza R. 1999. DnaSP, version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15:** 174–175.

Schluter D. 2001. Ecology and the origin of species. *Trends in Ecology & Evolution* **16:** 372–380.

Schoener TW. 1968. Anilis lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49:** 704–726.

Sites JW, Marshall JC. 2003. Delimiting species: a renaissance issue in systematic biology. *Trends in Ecology and Evolution* **18:** 462–470.

Sites JJ, Marshall JC. 2004. Operational criteria for delimiting species. *Annual Review of Ecology, Evolution and Systematics* **35:** 199–227.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* **15:** 4161–4293.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.

Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Systematic Biology* **57:** 758–771.

StatSoft, Inc. 2004. STATISTICA (data analysis software system), version 7. Available at: www.statsoft.com

Stockman AK, Bond JE. 2007. Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. *Molecular Ecology* **16:** 3374–3392.

Swets K. 1988. Measuring the accuracy of diagnostic systems. *Science* **240:** 1285–1293.

Toews DP, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21:** 3907–3930.

Van Valen L. 1976. Ecological species, multispecies, and oaks. *Taxon* **25:** 233–239.

Waltari E, Hijmans RJ, Peterson AT, Nyári AS, Perkins SL, Guralnick RP. 2007. Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE* **2:** e563.

Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62:** 2868–2883.

Wiens JJ, Graham CH. 2005. Niche conservatism: integrating evolution, ecology and conservation biology. *Annual Review of Ecology and Systematics* **36:** 519–539.

Zang DX, Hewitt GM. 2003. Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology and Evolution* **11:** 247–251.

Zink RM, Barrowclough GF. 2008. Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* **17:** 2107–2121.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Text S1. List of 18 morphometric variables and its respective acronyms, used in the morphometric analysis of *D. p. edwardsii*.

Table S2. Classification matrix after a discriminant analysis in *D. p. edwardsii*.

Figure S3. Tesselation map for population Southern lineage.

Figure S4. Tesselation map for Northern lineage.

Figure S5. Principal components plot for two molecular lineages of *D. p. edwardsii*, based on 18 morphometric variables.

Table S6. List of selected morphometric variables for a discriminant analysis in *D. p. edwardsii*, and obtained function coefficients for each variable considering lineage as groups.