doi:10.1111/evo.12924

Phylogeographic concordance factors quantify phylogeographic congruence among codistributed species in the *Sarracenia alata* pitcher plant system

Jordan D. Satler[1,2] and Bryan C. Carstens[1]

[1]*Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, Ohio 43210*

[2]*E-mail: satler.1@osu.edu*

Received July 13, 2015
Accepted April 1, 2016

Comparative phylogeographic investigations have identified congruent phylogeographic breaks in codistributed species in nearly every region of the world. The qualitative assessments of phylogeographic patterns traditionally used to identify such breaks, however, are limited because they rely on identifying monophyletic groups across species and do not account for coalescent stochasticity. Only long-standing phylogeographic breaks are likely to be obvious; many species could have had a concerted response to more recent landscape events, yet possess subtle signs of phylogeographic congruence because ancestral polymorphism has not completely sorted. Here, we introduce Phylogeographic Concordance Factors (PCFs), a novel method for quantifying phylogeographic congruence across species. We apply this method to the *Sarracenia alata* pitcher plant system, a carnivorous plant with a diverse array of commensal organisms. We explore whether a group of ecologically associated arthropods have codiversified with the host pitcher plant, and identify if there is a positive correlation between ecological interaction and PCFs. Results demonstrate that multiple arthropods share congruent phylogeographic breaks with *S. alata*, and provide evidence that the level of ecological association can be used to predict the degree of similarity in the phylogeographic pattern. This study outlines an approach for *quantifying* phylogeographic congruence, a central concept in biogeographic research.

KEY WORDS: **BUCKy, codiversification, community phylogeography, pleistocene, species tree distribution.**

Comparative phylogeography is considered to be a powerful, bottom-up approach that can leverage information from many single taxon studies to identify broad-scale population genetic trends across geographic space (Bowen et al. 2014). Although intraspecific studies may elucidate evolutionary processes and phylogeographic patterns, these results may be species-specific and not reflect overall evolutionary processes driving species and genetic diversity within a region (Page and Hughes 2014). Therefore, phylogeographic research depends on comparative investigations to fulfill its potential as an integrative discipline linking population genetics and phylogenetics (Avise et al. 1987), as similar results across multiple species likely result from shared responses to historical events (e.g., Arbogast and Kenagy 2001; Hickerson et al. 2010; Dawson 2013). In particular, Avise (2000) suggested that spatially congruent phylogeographic breaks among species be interpreted as shared responses to vicariance events, and the search for phylogeographic concordance has developed into a central objective of comparative phylogeography. The concept of phylogeographic concordance, however, has been entirely qualitative to date. Just as two phylogenetic trees can be either identical in one way or different in endless numbers of ways, phylogeographers have described species histories as either concordant or as not concordant, without any explicit measurement of the degree of discordance. This shortcoming limits the ability of the discipline to identify species that have responded in a concerted manner to large-scale changes in the environment, because monophyly (required before concordance can be recognized based on a visual inspection of genealogy) is dependent on the effective

population size (Ne) of the focal organism and requires substantial periods of time to form (Hudson and Coyne 2002). Furthermore, it is only identifiable with sufficient rates of nucleotide substitution, which may complicate comparisons between organisms with substantial differences in mutation rate, such as plants and arthropods.

Comparative phylogeographic investigations have typically focused on particular taxonomic groups that contain a number of codistributed taxa (e.g., Garrick et al. 2008; Bell et al. 2012; Fouquet et al. 2012; Oaks et al. 2012; Hope et al. 2014; Smith et al. 2014a), with a null model of species responding to historical events in an idiosyncratic manner. In these studies, shared phylogeographic breaks (sensu Swenson and Howard 2005) provide evidence that species responded to landscape-level changes in a similar way. Some comparative studies, however, have focused on ecological communities consisting of species from a variety of taxonomic groups (e.g., Whiteman et al. 2007; Roe et al. 2010; Espindola et al. 2014). Unlike other comparative phylogeographic studies, the expectation for investigations into ecological communities is for phylogeographic concordance across species, because the interacting members of an ecological community are to some extent codependent and thus should respond in a concerted manner to landscape changes. Phylogeographic investigations that document shared demographic responses to climatic and environmental events provide evidence for the stability of certain ecological interactions in evolutionary time (Smith et al. 2011). Although ecological interactions have long been thought to influence species diversification (e.g., Darwin 1859), the lack of a quantitative measure of concordance has prevented assessment of the degree to which ecological interactions and phylogeographic concordance are correlated.

A novel approach for measuring phylogeographic concordance is enabled by species tree phylogenetic methods (e.g., Edwards et al. 2007; Kubatko et al. 2009; Heled and Drummond 2010). We follow in historical biogeography (sensu Nelson and Platnick 1981) by aggregating samples from particular geographic regions into operational taxonomic units (OTUs). With such OTUs, concordance factors (Baum 2007) can be used to quantify the degree of phylogeographic concordance and to identify community-level patterns of diversification among the species. Concordance factors were originally introduced to infer a concordance tree (i.e., a species tree) from multiple gene trees, under the assumption that the most represented clade in the gene trees is also present in the species tree. Concordance factors do not require any assumptions about the evolutionary factors that cause incongruence among gene trees, nor do they explicitly estimate parameters such as population size or divergence times. As such, they represent an elegant approach for quantifying congruent genetic structure across codistributed species.

Concordance factors can be easily extended to comparative phylogeography, because recovered clades represent the proportion for which these relationships are true for a group of organisms. Phylogeographic concordance factors (PCFs) quantify one of the most important concepts in comparative phylogeography, concordance in the geography of phylogenetic breaks across co-occurring species (Type III concordance; Avise 2000), and secondarily provide an estimate of the dominant pattern of diversification from multiple, codistributed species. In the PCF framework, nodal support values (ranging from 0 to 1) designate the fraction of the community for which a particular clade is represented within the species' history. Communities exhibiting ecological relationships that are stable through time should contain higher phylogeographic concordance factors (as measured by PCF scores that approach 1), demonstrating a shared pattern of diversification. Our expectation is that tightly associated species will respond to landscape and historical events in a concerted manner, and this will be reflected in similar phylogenetic patterns. Alternatively, communities with species that exhibit an idiosyncratic response to historical events would be expected to contain lower phylogeographic concordance factors. In this article, we calculate phylogeographic concordance factors for an ecological community, conduct simulations to generate an expectation on the range of PCFs, and test the null hypothesis that ecological association is a predictor of phylogeographic congruence.

Ecological communities, defined here as an aggregate of codistributed species that are associated ecologically, contain some set of species that interact for at least part of their life cycle. These ecological interactions can be facultative or obligate—for example, species may be dependent upon other species for the acquisition of important nutrients, reproduction, or as hosts for a parasitic lifestyle. If these interactions are stable over long periods of time, species with obligate interactions should exhibit a high degree of phylogeographic concordance that may eventually manifest itself as a cophylogenetic pattern (Johnson and Stinchcombe 2007). For example, figs and fig wasps are obligate mutualists that share an intimate relationship whereby the insect is the sole pollinator of the plant. Phylogenetic data for both groups support a pattern of codiversification dating back 60 Ma (Rønsted et al. 2005). Host/parasite interactions can also be stable in evolutionary time, as best illustrated in pocket gophers and chewing lice (Hafner et al. 2003). Such studies demonstrate that ecological associations can extend into evolutionary time, and that complex processes that promote diversification can impact species in similar ways. If the evident cophylogeny in these ecologically dependent species reflects codiversification on the landscape scale, shared phylogeographic breaks on the landscape level should be common (e.g., Jousselin et al. 2008). By measuring the phylogeographic concordance at these early stages

(i.e., on the landscape level), the implicit prediction of such co-phylogenetic studies can be tested: that genetic concordance and the degree of ecological interaction are positively correlated.

An ecological community well suited for comparative phylogeographic research is the *Sarracenia alata* carnivorous pitcher plant system. *Sarracenia alata* inhabits longleaf pine savannahs and fens in the Gulf Coast of North America (McPherson 2007). Their leaves form as hollow, water-filled structures that enable the plant to capture and digest prey items, an adaptation for nutrient acquisition (Darwin 1875). These pitchers also comprise a unique ecosystem, where a broad diversity of associated organisms share complex interactions and relationships enabling the plant to act as host for a diversity of species (Adlassnig et al. 2011). Within *S. alata*, a diverse group of arthropods are known to interact ecologically with the plant, and can be categorized as obligate symbionts, herbivores, or capture interrupters (reviewed in Folkerts 1999). Obligate symbionts include several flesh fly species (family Sarcophagidae); these flies use the pitchers as an environment for larval deposit and development, with the plant providing both shelter and prey items during this life stage (Dahlem and Naczi 2006). Two described species of mites inhabit the pitchers; one is a scavenger in the pitcher fluid (*Sarraceniopus hughesi*), and the other feeds on nematodes and microscopic arthropod species (*Macroseius biscutatus*). The entire life cycle of the noctuid moth, *Exyra semicrocea*, occurs within the pitcher tubes, with the larval stage feeding upon the inner tissue of the leaf. A recent investigation into this moth utilizing mitochondrial DNA recovered a complete lack of haplotype sharing across the Mississippi River, and additional population structure east of the river (Stephens et al. 2011). In addition, multiple species of spiders act as capture interrupters, and opportunistically capture prey items that are attracted to the pitcher leaves. Although not restricted to the pine savannahs, the spiders live in close proximity to *S. alata* in these habitats.

Sarracenia alata and its arthropod community offer an ideal system for an exemplar community phylogeographic study. Population structure has been demonstrated at both shallow and deep levels within *S. alata* across its distribution (Zellmer et al. 2012; Carstens and Satler 2013), with congruent population genetic structure exhibited across multiple taxonomic groups recovered from the pitcher plant fluid (Satler et al. 2016). But apart from *E. semicrocea* (Stephens et al. 2011), little is known regarding the phylogeographic structure of the associated arthropod species. Here, we generate DNA sequence data from six arthropod species to investigate the evolutionary history of the *S. alata* pitcher plant system. We predict that ecological relationships correlate with phylogeographic congruence, use phylogeographic concordance factors to quantify the amount of phylogeographic congruence among members of the community, and identify species that do not share an evolutionary history with the pitcher plant system.

Material and Methods

TAXON SAMPLING

Samples from six arthropod species known to interact ecologically with *S. alata* were collected from throughout *S. alata's* distribution (Fig. 1). There were 12 total sampling sites; the sampling, however, varied across the locales (see Table S1 for sampling information). Arthropods included two species of flesh flies (*Fletcherimyia celarata*, *Sarcophaga sarraceniae*), one moth (*Exyra semicrocea*), one mite (*M. biscutatus*), and two spiders (*Misumenoides formosipes*, *Peucetia viridans*). All individuals were captured in or near the pitcher plant, with most resting on or just under the lid of the pitcher. For the two flesh fly species, males were pinned in the field with their genitalia extracted to confirm proper identification. For those fly specimens, three legs were removed and placed in 95% EtOH for DNA preservation; all other specimens were placed directly in 95% EtOH. Upon returning to the lab, all samples in ethanol were transferred to a –80° degree freezer for sample preservation. DNA was extracted from either crushed legs (flies, spiders) or full body soaking (moth, mite) using a Qiagen DNeasy kit; resulting samples were amplified for the mitochondrial cytochrome *c* oxidase subunit I (COI) barcode gene (Folmer et al. 1994) via polymerase chain reaction (PCR), with subsequent products sequenced in both directions. Sequences were edited using Sequencher v4.8 (Gene Codes Corporation, MI), and manually aligned using MacClade v4.08 (Maddison and Maddison 2005). Previously sequenced chloroplast and nuclear DNA (eight loci) was used for *S. alata* (Koopman and Carstens 2010; Zellmer et al. 2012; Carstens and Satler 2013).

POPULATION GENETIC STRUCTURE

Multiple summary statistics were generated to characterize genetic variation in the organellar genomes of the arthropod species. The packages ape (Paradis et al. 2004) and pegas (Paradis 2010) in the statistical platform R (R Development Core Team 2013) were used to calculate standard population genetic statistics, including π and Watterson's θ. G_{ST} values (Nei 1973) were calculated for each species in gstudio (Dyer 2012); these values measure the level of genetic partitioning among the sampling locales. As this takes into account haplotype data, an analysis of molecular variance (AMOVA; Excoffier et al. 1992) was estimated for each of the species to take full advantage of the sequence data. This was tested in a hierarchical manner, assessing genetic variation for (i) sampling locales within each side of the Mississippi River, (ii) sampling locales in the total distribution, and (iii) regions (i.e., east and west of the Mississippi River) within the total distribution. AMOVAs were calculated in the program SPADS (Dellicour and Mardulyn 2014), with 10,000 permutations to generate levels of significance. Spatial principal component analysis (sPCA)

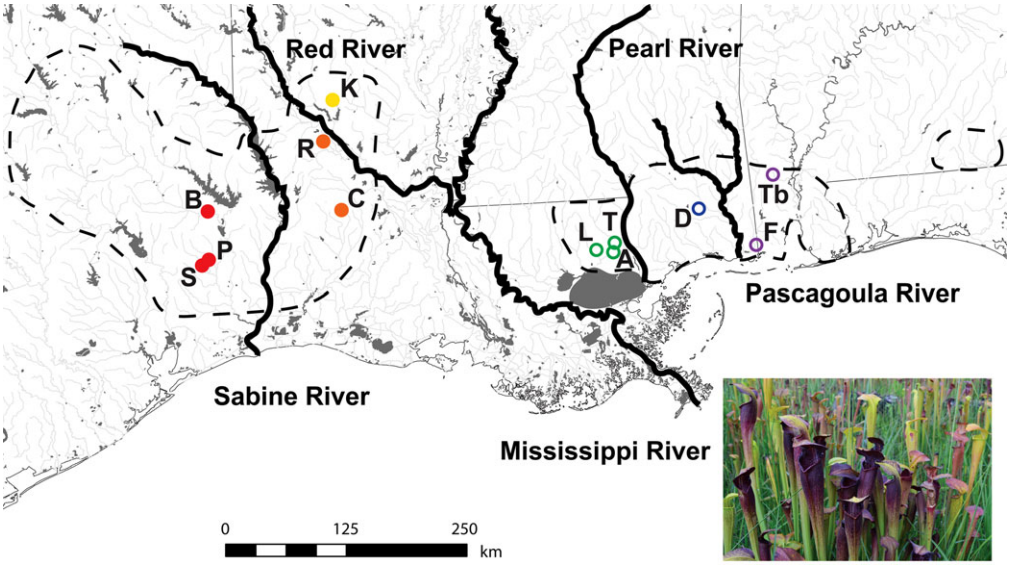

Figure 1. Distribution of *Sarracena alata* in the southeastern United States, with major rivers represented on the map. Broken lines indicate the known distribution of *S. alata*. Circles correspond to sampling sites for the study, with pattern of circles indicating the two hypothesized evolutionary lineages of the host plant (Carstens and Satler 2013). Locales are as follows: Sundew (S), Pitcher (P), Bouton Lake (B), Red Dirt (R), Cooter's Bog (C), Kisatchie (K), Lake Ramsey (L), Talisheek (T), Abita Springs (A), De Soto (D), Franklin Creek (F), and Tibbie (Tb). Taxa varied in the number of locales in which they are represented; details on the sampling distribution of each taxon can be found in Table S1. Insert picture is of *S. alata*.

in adegenet (Jombart 2008) and redundancy analysis (RDA) in vegan (Oksanen et al. 2015) were run to test how sampling location influenced genetic partitioning in the data, and to test if genetic breaks coincide with the Mississippi River. To test for a correlation between geographic and genetic distance, isolation by distance (IBD) values were calculated with the IBDWS v3.23 web server (Jensen et al. 2005). Geographic matrices were constructed measuring the Euclidean distance between sampling sites (in kilometers) with the distance measurement tool in Google Earth (www.google.com/earth/, last accessed Oct 27, 2015); genetic distance matrices for each species were generated using a Kimura 2–parameter (K2P) substitution model. In addition, maximum likelihood (ML) estimates of the gene tree topologies were estimated using Garli v2.01 (Zwickl 2006). Models of DNA sequence evolution were estimated in jModelTest v0.1 (Guindon and Gascuel 2003; Posada 2008). For each Garli analysis, datasets were partitioned by codon position, with 1000 bootstrap replicates to generate nodal support values. Support values were summarized on each tree with SumTrees v4.0.0 (Sukumaran and Holder 2015) from the python library DendroPy v4.0.3 (Sukumaran and Holder 2010).

PHYLOGEOGRAPHIC CONCORDANCE FACTORS

Phylogeographic concordance factors (PCFs) are calculated in a three-step process: (i) posterior distributions of species trees are estimated for each species independently, with all OTUs (i.e., tips in the trees representing geographic areas) represented across species, (ii) the frequency of each unique topology (relationships only) within each species tree distribution is summarized, and (iii) tree distribution summaries are combined across all species to generate a concordance tree. Nodal support values of the resulting concordance tree are the phylogeographic concordance factors, with the tree pattern representing the estimate of community diversification.

Sarracenia alata pitcher plant system

PCFs were calculated to test for a correlation between ecology and phylogeography, infer the dominant pattern of community diversification, and identify the species that share a congruent phylogeographic history with the host plant. Zellmer et al. (2012) proposed that habitat isolation facilitated by large rivers is the primary cause of population genetic structure in *S. alata*. This supports the idea that sampling locales within these regions were connected by suitable habitat prior to European settlement and subsequent land conversion and habitat loss (Noss et al. 1995). In order to calculate PCFs, we follow Zellmer et al. (2012) and define OTUs as the geographic regions divided by these major rivers. Specifically, sample locales were collapsed into the following five OTUs: Bouton Lake (B) + Sundew (S) + Pitcher (P); Cooter's Bog (C) + Red Dirt (R); Abita Springs (A) + Lake Ramsey (L) + Talisheek (T); De Soto (D); and Franklin Creek (F) + Tibbie (Tb) (see Fig. 1 for details); samples were

not included from Kisatchie (K) due to incomplete sampling of arthropod taxa. In addition, *M. formosipes* and *F. celarata* were not included as they did not contain sufficient sampling, leaving five species for the PCF analysis. Species tree distributions for each species were estimated with *BEAST v1.8.2 (Heled and Drummond 2010), associating alleles from each locale with their respective OTU. COI datasets were partitioned by codon position, with models of DNA sequence evolution as estimated above. Although single locus data are not ideal for species tree estimation, and we do not advocate this application absent a framework such as PCFs, mitochondrial DNA has been the marker of choice for phylogeography since the inception of the discipline, and in many cases does a reasonable job of tracking population history. Each individual analysis was performed under a strict molecular clock for 2.0×10^8 generations, 10,000 trees saved in the posterior distribution, and the first 10% of trees discarded as burn-in. For *S. alata*, an eight-locus dataset was used (one chloroplast and seven nuclear loci), including *Sarracenia rubra* as an outgroup for rooting purposes (data from Carstens and Satler 2013), and run for 5.0×10^8 generations (10,000 trees saved and 10% discarded as burn-in). Log files were imported into Tracer v1.6 (Rambaut et al. 2014) to check for convergence. Each species tree distribution was summarized using mbsum (Larget et al. 2010) to count the number of times each unique topology was represented within the posterior distribution. BUCKy (Ané et al. 2007) was then used to process the resulting tree summary files, generating a concordance tree with concordance factors. An alpha value of infinity was used, which represents a null hypothesis of no congruence. This approach is conservative, and thus requires strong support across two or more species before an inference of shared phylogeographic history is reached (based on PCF values). A custom python script (PCFs.py) was written to conduct steps ii and iii of the pipeline (https://github.com/jordansatler/PhylogeographicConcordance Factors).

If one or more species has not evolved with the community, it would be desirable to identify those taxa. To determine if a subset of taxa best fit a model of codiversification, all possible combinations of N taxa were analyzed using the pipeline, ranging from pairing just two species together to N–1 species. To compare the suite of models, average nodal support values (i.e., the average value of their community tree concordance factors) were calculated and partitioning schemes within each K level were ranked based on this value. Essentially, we would expect a dramatic increase in average PCF values if taxa that show discordant phylogeographic histories were removed from the analysis. Although there may be a hierarchical structure of expectations for PCF values (i.e., the deepest part of the trees are split by a biogeographic barrier, resulting in a high PCF value for that node, but there is relatively little substructure toward the tips in the trees, resulting in low PCF values for those nodes), averaging these values provides a way to quantitatively compare the various taxon partitioning schemes and identify potential taxa that do not fit within a community model.

Simulations for pairwise comparisons

Simulations were conducted to generate expectations of PCF values in pairwise comparisons ($K = 2$). 20 species trees were simulated under a Yule process with five OTUs and a total tree depth of 20N; trees varied randomly in regards to both interrelationships and branch lengths. Ten genealogies were simulated from each species tree with ms (Hudson 2002), and DNA sequence data were evolved from these genealogies using Seq-Gen (Rambaut and Grass 1997) under an HKY model. After generating the sequence data, each simulated DNA matrix was treated as an individual (analogous to the COI data from each arthropod), and species tree distributions were estimated using *BEAST following conditions stated above. In total, the analysis produced 200 species tree distributions, with ten distributions per species tree topology. PCFs were calculated in a pairwise manner (i.e., the $K = 2$ level), testing taxon pairs that were generated from the same species tree. Average PCF values were summarized to generate a distribution of expected values when species shared the same phylogeographic history. Simulations therefore take into account coalescent stochasticity and species tree estimation error.

Proof of concept in other systems

We applied the PCF method to two additional systems, varying in ecological associations among the included species. We analyzed Sanger sequence data described in Carstens et al. (2005) from the Pacific Northwest (PNW) temperate rainforest, and ultraconserved DNA element (UCE) data described by Smith et al. (2014b) from Neotropical birds. The former included three amphibians (one frog and two salamanders), a small mammal (vole), and a tree (willow) distributed across the Cascade and Rocky Mountains (see Carstens et al. 2005 for details). Species tree distributions were estimated with *BEAST (as described above) from a combination of single locus (frog, *Dicamptodon*, vole, willow) and multilocus (*Plethodon*) datasets, with four geographic regions constituting the OTUs. The latter included next-generation sequence data from Smith et al. (2014b), consisting of four bird species spanning three biogeographic barriers in South and Central America (see Smith et al. 2014b for sampling details). We included four of the five species from the investigation, as the fifth species was not sampled from all geographic regions. *Schiffornis turdina* contained one region that Smith et al. (2014b) divided into two, but since these two regions formed a clade with $pp = 1.0$, we collapsed the node and treated the two areas as a single

region in our analysis. Species tree distributions were provided by B. Smith (those analyzed in Smith et al. 2014b); each species tree distribution was estimated with between 115 and 144 UCE loci (varying by species). PCF analyses for both systems were run as outlined above.

TESTS OF SIMULTANEOUS DIVERGENCE IN *S. alata* SYSTEM

The msBayes model (Hickerson et al. 2006; Hickerson et al. 2007; Huang et al. 2011) was used to test for simultaneous divergence across the Mississippi River among the community members. In contrast to PCFs, msBayes evaluates the timing of divergence among population pairs to understand if a model of simultaneous divergence best explains the data. Given the complexity of the model, a full likelihood approach is computationally intractable and hierarchical Approximate Bayesian Computation (hABC) is used. The msBayes method estimates hyperparameters that are of interest to the researcher, including posterior probabilities on the number of divergence episodes, mean, and variance of the timing of the divergence episodes. An appealing aspect of this method is it takes into account both mutational rate variation and coalescent stochasticity within each species, so lineage-specific aspects of their molecular evolution are accounted for when estimating if the timing of divergence is synchronous.

We used an isolation only model where the Mississippi River divides the community and migration ceases following this split. The Mississippi River is a well-characterized biogeographic barrier (reviewed in Soltis et al. 2006), with a lack of suitable habitat for *S. alata* in the Atchafalaya Basin leading to a disjunct distribution in the plant. Research on *S. alata* shows the deepest split in the population tree corresponding with the Mississippi River (Zellmer et al. 2012), with populations restricted to either side of the Mississippi River representing evolutionary lineages (Carstens and Satler 2013). In addition, four of the six arthropods show reciprocal monophyly in their COI genealogies, as well as population genetic structure in spatial analyses corresponding with the Mississippi River (see Results).

We applied the updated hABC model implemented in PyMsBayes (Oaks 2014), with all species (*rps*16–*trn*K gene used for *S. alata*) included in the analysis. PyMsBayes has multiple changes that allow for improved accuracy with the msBayes model; primarily, rather than placing a uniform prior distribution on the number of divergence episodes, PyMsBayes places a Dirichlet process prior on all of the divergence models, with gamma prior distributions used on many parameters, including population divergence (τ) and population size (θ). The divergence episode prior distribution was centered around three divergence episodes, as this represents our prior hypothesis of diversification; the two spider species are characterized as capture interrupters, and do not show any obligate association with the plant. In contrast, given the life-history traits of the other arthropod species, we may expect the ecologically associated species to show similar divergence patterns to *S. alata.* Population sizes parameters were either estimated individually (ancestral and both daughter populations) or as one (all populations given the same population size parameter). Reciprocal monophyly in a single locus dataset creates difficulty in estimating both timing of population divergence and ancestral population size, as the lack of incompletely sorted alleles leaves little information to the size of the ancestral population (Edwards and Beerli 2000). Given that four of the seven species do show this pattern (see Results), difficulty in estimating ancestral population size could lead to inaccuracy in the estimation of divergence times; restricting population sizes to a single parameter may help combat this issue. The prior distribution on divergence times had an upper limit at the Pleistocene/Pliocene change, with much of that distribution less than 1 Ma, consistent with divergence estimates of *S. alata* (Zellmer et al. 2012). As *S. alata* is the host for the ecological community, we informed prior choices based on our knowledge of the evolutionary history of the plant. Generation times and mutation rates were modeled independently in each species; specific details on the PyMsBayes configs files can be found in Supplemental Material. Analyses were run for 3×10^6 generations, with 1000 samples drawn to generate the posterior distribution. The simulated and empirical data were compared using the default combination of summary statistics, with the unadjusted posterior estimate used to represent the posterior distribution (Wegmann et al. 2010). The vector of the summary statistics was not reordered, preserving the original taxon order of the observed data as suggested by Oaks (2014).

Results

GENETIC SAMPLING

Between 21 and 138 COI mtDNA sequences were generated for each arthropod species (Table 1). All newly generated sequences have been deposited in GenBank under accession numbers KU975801–KU976138 (see Table S1 for more details). Sequences from the *rps*16–*trn*K chloroplast gene (Koopman and Carstens 2010) and seven anonymous nuclear loci (Zellmer et al. 2012; Carstens and Satler 2013) were gathered from previous studies for the analysis of *S. alata*. Datasets and input files have been deposited in Dryad (XXXX). Q

POPULATION GENETIC STRUCTURE

Genetic diversity varied across sampled taxa (Table 1). Some species exhibited high numbers of unique haplotypes (e.g., *M. biscutatus*, 13 out of 21), while others had low haplotype diversity (e.g., *F. celarata*, 4 out of 33). Nucleotide diversity ranged

Table 1. Sampling and sequencing information for all arthropod species.

Species	N	H	SS	π	θ_w	Tajima's *D*	G_{ST}
Sarcophaga sarraceniae	157	13	16	0.0059	0.0046	0.7060	0.4774*
Fletcherimyia celarata	33	4	7	0.0028	0.0028	0.0105	0.5859*
Exyra semicrocea	37	19	31	0.0142	0.0117	0.7506	0.4453*
Macroseius biscutatus	21	13	71	0.0515	0.0309	2.4239*	0.6146*
Peucetia viridans	56	18	28	0.0041	0.0096	–1.8789	0.0834
Misumenoides formosipes	34	28	38	0.0086	0.0149	–1.5251	0.2927

Information includes the number of sequences (N) and unique haplotypes (H) per species, segregating sites (SS), nucleotide diversity (π), Watterson's θ per site (θ_w), Tajima's *D*, and G_{ST} values. Asterisks indicate significant *P*-values below 0.05.

Table 2. AMOVA and IBD results.

	AMOVA			IBD		
Species	Φ_{SC}	Φ_{ST}	Φ_{CT}	*r*	RMA slope	r^2
Sarcophaga sarraceniae	0.0117	0.9362*	0.9355*	0.8635*	0.00319	0.746
Fletcherimyia celarata	–0.1879	0.9690*	0.9739	0.8087	0.00564	0.654
Exyra semicrocea	0.4548*	0.9059*	0.8274*	0.6782*	0.00242	0.460
Macroseius biscutatus	0.4911*	0.9782*	0.9571*	0.6034*	0.00258	0.364
Peucetia viridans	–0.0148	–0.0125	0.0023	0.0478	0.00078	0.003
Misumenoides formosipes	–0.0406	–0.0427	–0.0021	NA	NA	NA

Asterisks indicate significant *P*-values below 0.01. Isolation by distance was not calculated in *M. formosipes* due to limited sampling in multiple localities.

from about 0.003 to 0.05, and Watterson's θ ranged from 0.003 to 0.01. Values of Tajima's *D* varied from –1.88 to 2.42, with one taxon containing a significant *D* statistic (*M. biscutatus*, 2.42). Population structure is evident in multiple taxa within the dataset. G_{ST} values are high, with an average of 0.478 (Table 1). Except for the spider species, G_{ST} values from all species are statistically significant, with values greater than 0.445. Results from the AMOVAs (Table 2) are similar, with significant genetic structuring in a minimum of one level for all species with the exception of the two spiders, where no structure was detected; the moth and mite show significant population structure at all three levels. Population structure is also recovered in the sPCA (Fig. S1) and RDA analyses (Table S2) for the nonspider arthropods. Limitations in dispersal ability and ecological associations with the host plant are likely contributing to the high levels of population genetic structure. This pattern is further reflected in the COI gene tree estimates (Figs. S2–S7). Reciprocal monophyly is detected in all species minus the spiders, and highlights the Mississippi River as a biogeographic barrier influencing genetic variation in many of these species. Strong support values are seen in the trees, although finer scale structure is less present in some of the taxa. Overall, the spiders show essentially no population genetic structure, a result that may reflect their ability to disperse over large distances (e.g., Greenstone et al. 1987).

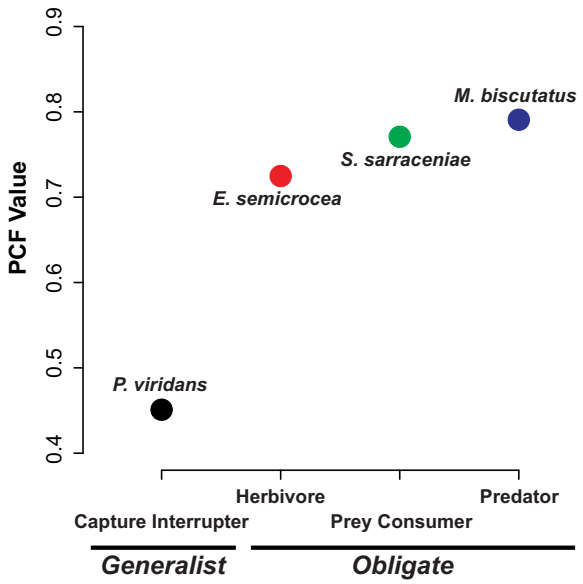

Figure 2. Phylogeographic concordance factors between each arthropod species and the host pitcher plant. Pairwise comparisons represent the amount of congruence between the species' posterior distributions of species trees (see Table 3 for specific PCF values). Higher host-plant association correlates with a higher average phylogeographic concordance factor. Ecological categorical types from Folkerts (1999).

COLOUR ONLINE, B&W IN PRINT

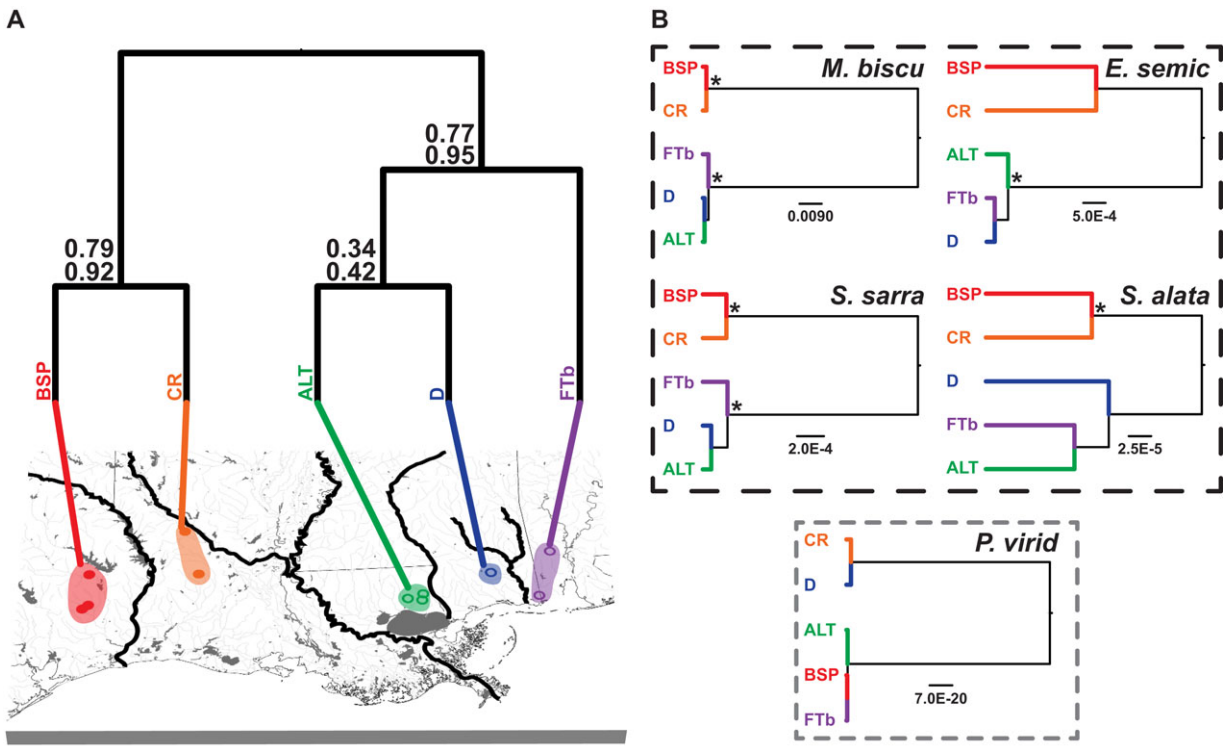

Figure 3. **Community pattern of diversification estimated with phylogeographic concordance factors. Panel A represents the community tree estimated in BUCKy. Concordance factors on top reflect when all five species are used; concordance factors on bottom are with *Peucetia viridans* removed from the PCF analysis. Panel B represents the maximum clade credibility (MCC) trees of the *BEAST results from each species, providing a visual perspective of the two recovered groups. Posterior probability values greater than 0.95 are represented by an asterisk. Species names have been truncated for illustration purposes.**

PHYLOGEOGRAPHIC CONCORDANCE FACTORS

Results from S. alata system

Pairwise comparisons between each arthropod species and the host plant demonstrate that ecological association is a good predictor of phylogeographic congruence (Fig. 2). In particular, the three species known to interact ecologically with *S. alata* have average PCF values above 0.71 (highest between the mite and host plant at 0.79), while the average PCF value for the spider and plant is 0.45. To assess the significance of these values, we analyzed simulated data from 900 pairwise comparisons for taxa where each species pair was from the same model of diversification (i.e., the same species tree); these results provided a benchmark for interpreting the PCF values calculated from the empirical data (Fig. S8). The 90% highest posterior density interval for PCFs (90% HPD = 0.71–0.99) suggests that the values observed in the species that interact ecologically with the plant are indicative of codiversification. The value for the spider, on the other hand, is sufficiently low as to suggest no shared history of diversification. This result holds when the 95% highest posterior density, a more inclusive distribution, is used (95% HPD = 0.62–0.99).

When PCFs are calculated from sets of more than two species, they provide evidence for a shared phylogeographic history among the members of the community. With all five species included, fairly strong support emerges for a shared east/west split, with concordance factors of 0.79 for a western clade and 0.77 for an eastern clade (Fig. 3A). A concordance factor of 0.34 shows less support for a clade uniting ALT and D, suggesting that phylogeographic relationships within the eastern region are less concordant across the species. Results from conducting all permutations of taxa at all K levels (greater than one) provide evidence that *P. viridans* has not codiversified with the community (Table 3). For example, when all permutations are run at $K = 4$ (dropping one species each time), average nodal support values are substantially higher when the spider is not included. At the $K = 3$ level, all permutations not including *P. viridans* again show a substantially higher PCF average than any community compositions including the spider. This pattern is also evident at the $K = 2$ level, indicating that *P. viridans* has not codiversified with the community. Removing the outlier species from the analysis returns the same community tree (Fig. 3A), but increases nodal support values for the western (to 0.92) and eastern (to 0.95) clades. As further substructure in the east is not well supported (increase to 0.42), PCFs suggest that response to the Mississippi River has driven the shared phylogeographic patterns seen in this

community. The pattern can be visualized with the maximum clade credibility (MCC) trees from each of the species, where two distinct groups are seen (Fig. 3B).

Results from other systems

We analyzed comparative data sets in two additional systems with the PCF framework. In the PNW, results are consistent with the basic patterns described by Carstens et al. (2005), and reinforce previous assertions of phylogeographic congruence (PCF ≥ 0.98) among the three amphibian species (Fig. 4A). PCFs identify the vole as an outlier to the community-level pattern, as pairwise comparisons that include the vole have low average PCF values (∼0.5; Fig. 4B). PCF values also decrease when the willow tree is included with the amphibians (Table S3). In the Neotropical bird dataset, results show generally high levels of phylogeographic congruence among the species (Fig. 4C; Table S4), with all pairwise comparisons above 0.75 (Fig. 4D). This demonstrates relatively similar topological patterns among the species (see Smith et al. 2014b for divergence time comparisons among species).

TESTS OF SIMULTANEOUS DIVERGENCE IN *S. alata* SYSTEM

Results from PyMsBayes suggest there have been multiple divergence episodes in this system (Table 4). The posterior distribution, however, is similar to the prior distribution, suggesting a lack of information present in the data. These results do not appreciably change when effective population sizes are treated as equal. To further explore the sensitivity of the analyses to the priors, the prior distribution on divergence episodes was adjusted to reflect a community where nearly all species diverged idiosyncratically (centered around six divergence episodes). Results from these analyses—one containing a different population size for each population, the other with the same parameter for all three populations—show posterior distributions once again similar to the prior distributions (Table S5). While there is little to no support for a single episode of divergence, it appears the data are insufficient for such a parameter-rich model.

Discussion

PHYLOGEOGRAPHIC CONCORDANCE FACTORS ALLOW CONGRUENCE AMONG TAXA TO BE QUANTIFIED

Phylogeographic concordance factors allow the degree of phylogeographic congruence to be quantified across multiple species, enabling researchers to identify species that share a pattern of codiversification. To date, comparative phylogeographic studies have commonly compared phylogenies across multiple taxa in a qualitative manner, and interpreted similarity as a shared response to a historic event (Avise 2000). However, to paraphrase Tolstoy, phylogenies that are not identical are each incongruent in their own unique way. So long as researchers rely on qualitative interpretations of congruence, comparative phylogeography will be prone to overinterpretation (Knowles and Maddison 2002).

PCFs provide two important contributions to comparative phylogeography. First, PCFs allow researchers to *quantify* the amount of congruence among the phylogeographic histories of codistributed species. This method can be applied to codistributed species within certain taxonomic groups (e.g., birds, reptiles, mammals), as has been historically done in comparative phylogeographic research, and can help identify earth history processes that have promoted diversification and structured genetic and taxonomic variation in similar ways. When applied to an ecological community, however, they allow for an assessment of the correlation between ecological association and evolutionary history. Although multiple macroevolutionary processes can influence phylogenetic congruence (e.g., dispersal, selection, speciation,

Table 3. Phylogeographic concordance factors for all permutations of *K* levels from two to N total species.

K	$PCF_{average}$	Species composition
5	0.63	*M. biscu, E. semic, P. virid, S. alata, S. sarra*
4	0.77	*M. biscu, E. semic, S. alata, S. sarra*
4	0.62	*M. biscu, E. semic, P. virid, S. sarra*
4	0.62	*M. biscu, P. virid, S. alata, S. sarra*
4	0.59	*M. biscu, E. semic, P. virid, S. alata*
4	0.59	*E. semic, P. virid, S. alata, S. sarra*
3	0.80	*M. biscu, E. semic, S. sarra*
3	0.79	*M. biscu, S. alata, S. sarra*
3	0.75	*M. biscu, E. semic, S. alata*
3	0.73	*E. semic, S. alata, S. sarra*
3	0.61	*M. biscu, P. virid, S. sarra*
3	0.57	*M. biscu, E. semic, P. virid*
3	0.56	*M. biscu, P. virid, S. alata*
3	0.56	*P. virid, S. alata, S. sarra*
3	0.55	*E. semic, P. virid, S. sarra*
3	0.54	*E. semic, P. virid, S. alata*
2	0.85	*M. biscu, S. sarra*
2	0.80	*M. biscu, E. semic*
2	0.79	*M. biscu, S. alata*
2	0.77	*S. alata, S. sarra*
2	0.74	*E. semic, S. sarra*
2	0.72	*E. semic, S. alata*
2	0.51	*M. biscu, P. virid*
2	0.47	*P. virid, S. sarra*
2	0.45	*P. virid, S. alata*
2	0.45	*E. semic, P. virid*

Average PCF values are calculated by taking the average of the concordance values at all ingroup nodes. For each *K* level, the taxonomic compositions are sorted to show the species groupings with the highest value of phylogeographic concordance. Taxon names have been truncated.

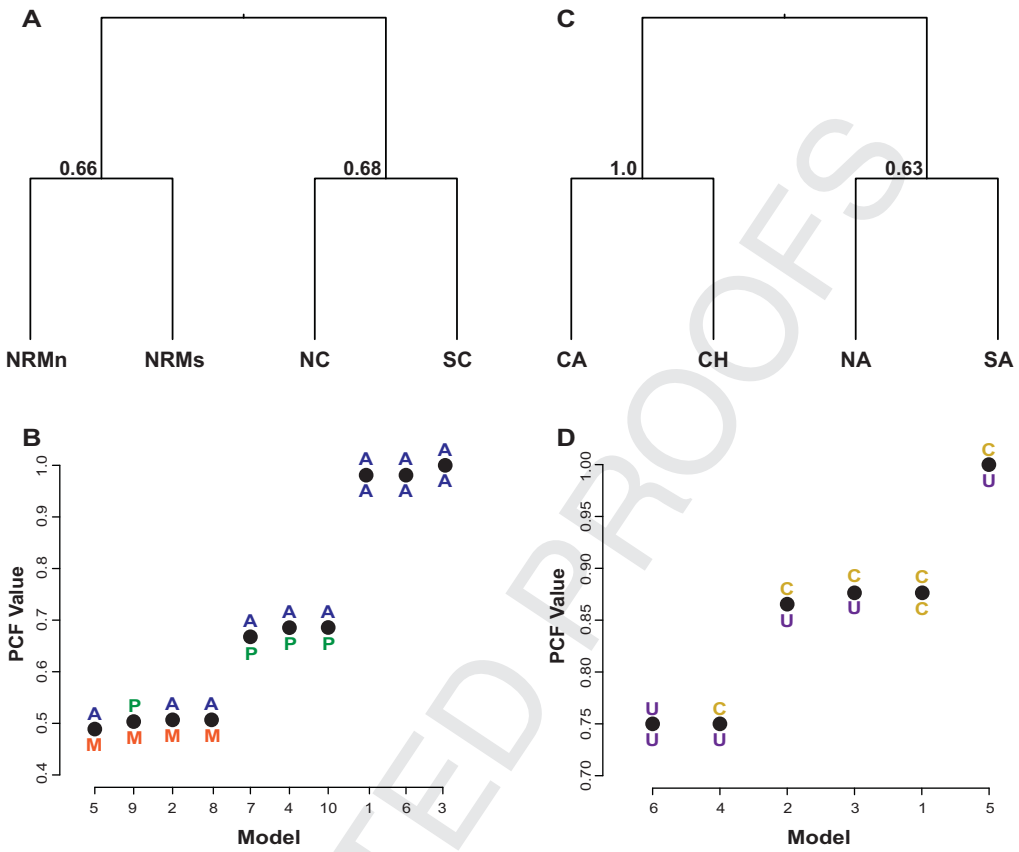

Figure 4. Phylogeographic concordance factors estimated in two additional data sets: Pacific Northwest and South and Central America. Panels A and B are results from Carstens et al. (2005); Panels C and D are results from Smith et al. (2014b). Panel A shows the community tree for the PNW, with the deepest split between the Cascade and Rocky Mountains, and moderate support at both nodes. Panel B shows all pairwise comparisons; all models containing two amphibians (A) show essentially identical phylogeographic patterns, while models containing the small mammal (M) reflect high levels of discordance. Panel C shows all four bird species contain a CA, CH sister relationship, with less support for an NA, SA relationship. Panel D shows all pairwise comparisons have a minimum average PCF value of 0.75, reflecting relatively high concordance of topological relationships among all bird species. Birds are categorized as living in the canopy (C) or understory (U); see Smith et al. (2014b) for sampling details.

and extinction), a null prediction is that ecological dependence is correlated with phylogenetic congruence (Clayton et al. 2004). While this has been best demonstrated in host/parasite (Hafner et al. 2003) and plant/arthropod interactions (Farrell and Mitter 1990; McKenna et al. 2009), ecological communities that consist of many mutualist species represent potentially informative systems for phylogeographic investigation. In the *S. alata* system, arthropod species that have a tighter ecological association with the pitcher plant exhibit higher average PCF values, a pattern predicted based on ecology (Fig. 2). Second, PCFs allow the researcher to potentially identify taxa that do not fit the overall pattern of community diversification. If evolutionary processes have strongly influenced diversification in a region, or the community contains tight ecological associations, it may be reasonable to expect that taxa share a phylogeographic history. By comparing different models of community composition, PCFs allow for the identification of groups of organisms that have responded in similar ways to historical events, and in the case of those species that share ecological association, to identify communities with taxa that have sustained their ecological relationships through evolutionary time. For an example of the latter, Table 3 shows the average phylogeographic concordance factors across 26 models, varying in taxonomic composition. All models that exclude *P. viridans* contain PCF values within the 90% HPD of expected values (see Fig. S8), but models that include *P. viridans* are not contained within the range of values expected if the species share a phylogeographic history (based on our simulations at the $K = 2$ level). This compelling result corroborates the idea that the spider is an opportunistic predator taking advantage of this specialized habitat when present, but does not rely on this habitat for survival and reproduction.

Phylogeographic concordance factors provide a flexible approach for comparative phylogeography because they are calculated directly from species tree distributions. PCFs can be

Table 4. Results from PyMsBayes analyses.

Divergence episodes	Unique models	Prior	$\theta_A \neq \theta_D$ Posterior	$\theta_A = \theta_D$ Posterior
1	1	0.1170	0.1190	0.1550
2	63	0.2710	0.2910	0.3100
3	301	0.3098	0.2860	0.2930
4	350	0.2020	0.2030	0.1730
5	140	0.0776	0.0770	0.0610
6	21	0.0202	0.0210	0.0050
7	1	0.0024	0.0030	0.0030

Prior distribution is centered around three divergence episodes, consistent with our prior beliefs on the system. Posterior distributions are very similar to the prior distribution, regardless of how population sizes (ancestral and daughter) are treated.

calculated from single or multilocus datasets (as demonstrated here), based on either full sequence data or single nucleotide polymorphisms (SNPs). While we analyzed single locus data for the arthropod species, multilocus data are preferable for PCF analysis, as data gathered from multiple independent nuclear markers will improve the estimate of phylogenetic history (Rokas and Carroll 2005). Single locus data have recognized shortcomings for phylogeography, particularly for parameter estimation (Edwards and Beerli 2000; Felsenstein 2006), but the COI gene (used here) represents an informative marker (e.g., Herbert et al. 2003). Multilocus data (eight loci) were analyzed for the host plant, and the PCF framework allows for the quantitative integration of analyses using both types of data. The use of the species tree framework also greatly reduces the impact of phylogenetic error that results from the stochastic process of allele coalescence, a necessary requirement as most phylogeographic research is focused in recent time scales (e.g., Pleistocene Epoch). The Bayesian inference framework accounts for phylogenetic uncertainty; by utilizing a posterior distribution of trees, error inherent to point estimates is thus avoided. The reliance on posterior distributions of trees, however, should be taken into account when interpreting the absolute values of the PCFs. An average PCF value of 1.0 (i.e., perfect phylogeographic concordance) can only be achieved if all trees in the posterior distributions across all species are topologically identical. As this is highly unlikely to ever be the case, species are unlikely to reflect average PCF values of 1.0, even if they have codiversified. For example, our results when testing simulated data sets that came from the same tree highlight this issue, as an average PCF value of 1.0 was never recovered (max = 0.9973). The median PCF average value under these conditions was 0.86, reflecting high phylogeographic concordance, with values down to 0.71 still found within the 90% HPD (even though they were simulated from the same tree). Depending upon the process of diversification and speciation in the taxa, expectations may vary and a range of values could be indicative of a pattern of codiversification. As the simulation conditions used here matched the empirical data, increasing the number of loci will improve the phylogenetic estimate of the species, and will likely increase the average PCF values for codistributed taxa if they do share a pattern of codiversification.

To demonstrate the efficacy and flexibility of phylogeographic concordance factors, we applied the method to two additional systems. The three datasets span the ecological continuum, from species within the same taxonomic group (Neotropical birds), to species from disparate taxonomic groups with some sharing habitat requirements (PNW), to species from disparate taxonomic groups with some tightly linked through ecology (*S. alata*). The pitcher plant system may be thought of as a best case scenario, where strong ecological interactions have manifested themselves into shared phylogeographic patterns; however, high PCF values are also seen in the other two systems. As these additional datasets contrast the pitcher plant system (in their ecological interactions), they allow us to tease apart how evolutionary and ecological processes contribute to species interactions over historical time. For example, habitat requirement explains why the three amphibians (in the PNW) share nearly identical phylogeographic histories, as there is no evidence that they are ecologically dependent on one another. In total, these analyses demonstrate that PCFs are broadly applicable to a diverse range of phylogeographic data. Because they are broadly applicable, they provide an important tool for synthesis between comparative phylogeography and ecology.

THE MISSISSIPPI RIVER IS A DRIVER OF DIVERSIFICATION

The Mississippi River is a well-characterized biogeographic barrier (Soltis et al. 2006), and has had a strong effect on the *S. alata* system, restricting gene flow across this large body of water for many taxa. Strong population genetic structure is seen among most arthropod species (G_{ST} values and AMOVA results; see Tables 1 and 2), suggesting that the landscape processes that promote diversification within *S. alata* have a similar influence on the associated arthropods. The PCF analysis provides strong evidence in support of this finding. Shared phylogeographic patterns

are seen with the fly (*S. sarraceniae*), the mite, and the moth, as compared to the host plant (Fig. 2), as well as when estimating a community phylogeny (Fig. 3A). While the reciprocal monophyly observed in the gene trees for four of the six arthropods further supports the isolation of populations on either side of this biogeographic barrier, the PCF analysis indicates that all of the major rivers that bisect the range of *S. alata* play a similar role in inhibiting gene flow. In addition, the PCF analysis complements other comparative phylogeographic approaches, such as hABC tests of simultaneous divergence.

We used PyMsBayes to test for simultaneous divergence across the Mississippi River. Unlike PCFs, which makes inference based on topology only, PyMsBayes estimates divergence times for all taxa and estimates the posterior probability of the number of divergence episodes for a given system. Implementation of ABC analyses is challenging, as multiple steps in the process can lead to incorrect inference, including the selection of summary statistics (Robert et al. 2011) and models (Pelletier and Carstens 2014) to include in the analysis. Results from analysis of the *S. alata* system with PyMsBayes were inconclusive, as the posterior distribution aligned closely to the prior distribution. Although the model was informed by species-specific information from the literature (see Supplemental Material), it is likely that patterns in the mtDNA are driving these results, specifically that four of the six arthropod species are reciprocally monophyletic across the Mississippi River in their ML tree estimates (Figs. S2–S7); *S. alata* is also monophyletic in the eastern region of the chloroplast ML tree (Fig. S9). Estimating population divergence times is exceedingly difficult under this scenario, as different combinations of ancestral population size and gene divergence time can lead to the same population splitting event. The confounding of these two parameters (with single locus data) leads to a large variance in divergence time estimation, necessarily making this approach uninformative in our study, regardless of how we designed the analysis (Table 4; Tables S5–S6). It appears that the data collected here, when summarized following Hickerson et al. (2006), contain little information as to the number of divergence events in this community when estimating temporal synchronicity of population splitting times in a parameter-rich model.

PCFs are strictly a spatial test and do not explicitly account for the timing of divergence. In combination with the results of the hABC analysis (when taken at face value), the taxa that share the same phylogeographic break in the pitcher plant system may have diversified at different times, although a lack of signal in the data highlight the difficulty in robustly estimating times of divergence (Arbogast et al. 2002). Since the biggest shortcoming of the single locus data is their inability to estimate population divergence times (Edwards and Beerli 2000), multilocus data are likely to improve our ability to test for codiversification, where shared spatial patterns can be further assessed for temporally concordant divergences. A full analysis can be seen with the Neotropical birds, where inferences based on topology (with PCFs) complement, but somewhat vary, with inferences based on divergence times (see Smith et al. 2014b). PCFs provide a quantitative tool for the comparison of topological patterns; we envision the addition of this novel method to the phylogeographer's toolbox, where analysis of divergence times and population sizes (in addition to PCFs) can provide a full understanding of diversification dynamics in codistributed species.

INSIGHTS INTO THE *S. alata* PITCHER PLANT SYSTEM

Population structure has been demonstrated in *S. alata* at both shallow and deep levels in the population tree (Koopman and Carstens 2010; Zellmer et al. 2012; Carstens and Satler 2013). Given the arthropod community that interacts ecologically with the plant (Folkerts 1999), ecological association predicts phylogeographic concordance among the community members. Congruent population structure is recovered across the fly, moth, and mite, which suggests that the evolutionary processes shaping genetic variation in the Gulf Coast have had a similar effect on those species as on the host plant. All three of these arthropods have high affinity for the pitcher plant habitat: *E. semicrocea* spends its entire life cycle within the plant, only moving around at night for short-range dispersal and mating, and both the moth and *S. sarraceniae* (Dahlem and Naczi 2006) depend upon the pitcher for larval deposition, with their larvae feeding and developing within the pitcher tubes. The mite largely preys upon nematodes and smaller mites within *S. alata*'s pitcher, and is seldom collected outside of this environment (Muma and Denmark 1967). All of these species have a high degree of phylogeographic concordance (Fig. 3)

In contrast, there is little evidence that the spiders are congruent with *S. alata*. Both spiders are opportunistic predators and widespread species that are not ecologically dependent on *S. alata*. Both also exhibit ballooning (Foelix 1982), a mechanism for long-range dispersal, and this behavior, in combination with their lack of ecological dependency with *S. alata*, likely explain the lack of concordance demonstrated here.

Conclusions

The *S. alata* pitcher plant system exhibits, at the broadest scale, shared phylogeographic patterns of multiple arthropods suggestive of codiversification. Ecological relationships within the system predict the level of phylogeographic congruence among the species, demonstrating that the flies, moth, and mite have a shared phylogeographic history with the host plant. The spiders, which are not ecologically dependent on *S. alata* and are capable dispersers, show no evidence of codiversifiication. To better understand the history of this community, we introduced

phylogeographic concordance factors. PCFs quantify phylogeographic congruence among species, whether in pairwise comparisons between individual species and the host plant (Fig. 2), among all species (Fig. 3A), or permutations of all combinations of species at varying *K* levels (Table 3). This novel approach can utilize the posterior distribution from any Bayesian species tree method, and is applicable to differing amounts and types of phylogeographic data, including full sequences from any number of loci as well as SNPs (through the use of a program-like SNAPP; Bryant et al. 2012). This inherent flexibility enables PCFs to be applied to data collected from thousands of previously published phylogeographic studies, making a valuable contribution to our understanding of how evolutionary processes have shaped and structured taxonomic and genetic variation within a region.

ACKNOWLEDGMENT

We thank Meg Daly, Ben Titus, Paul Blischak, and members of the Carstens lab for comments and discussion regarding the manuscript. We thank Rob Naczi and Greg Dahlem for assistance with identifying fly specimens, and Brian Tilston Smith for sharing species tree distributions from Smith et al. (2014b). We also thank A.E. Burbrink and three anonymous reviewers for comments that helped to improve this manuscript. Funding was provided by the National Science Foundation Doctoral Dissertation Improvement Grant (DEB-1501474), the Society for the Study of Evolution (Rosemary Grant Award for Graduate Student Research), and the Society of Systematic Biologists (Graduate Student Research Award) to J.D.S., and the Louisiana EPSCoR Pfund, the LSU Office of Research and Economic Development and Louisiana Board of Regents to B.C.C.

DATA ARCHIVING

The doi for our data is https://github.com/jordansatler/PhylogeographicConcordanceFactors.

LITERATURE CITED

Adlassnig, W., M. Peroutka, and T. Lendl. 2011. Traps of carnivorous pitcher plants as a habitat: composition of the fluid, biodiversity and mutualistic activities. Ann. Bot. 107:181–194.

Ané, C., B. Larget, D. A. Baum, S. D. Smith, and A. Rokas. 2007. Bayesian estimation of concordance among gene trees. Mol. Biol. Evol. 24:412–426.

Arbogast, B. S., and G. J. Kenagy. 2001. Comparative phylogeography as an integrative approach to historical biogeography. J. Biogeogr. 28:819–825.

Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annu. Rev. Ecol. Syst. 33:707–740.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18:489–522.

Avise, J.C. 2000. Phylogeography: the history and formation of species. Harvard Univ. Press, Cambridge, MA.

Baum, D. 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. Taxon 56:417–426.

Bell, R. C., J. B. MacKenzie, M. J. Hickerson, K. L. Chavarria, M. Cunningham, S. Williams, and C. Moritz. 2012. Comparative multi-locus phylogeography confirms multiple vicariance events in co-distributed rainforest frogs. Proc. R. Soc. B. 279:991–999.

Bowen, B. W., K. Shanker, N. Yasuda, M. C. D. Malay, S. von der Heyden, G. Paulay, L. A. Rocha, K. A. Selkoe, P. H. Barber, S. T. Williams, et al. 2014. Phylogeography unplugged: comparative surveys in the genomic era. Bull. Mar. Sci. 90:13–46.

Bryant, D., R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29:1917–1932.

Carstens, B. C., S. J. Brunsfeld, J. R. Demboski, J. M. Good, and J. Sullivan. 2005. Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. Evolution 59:1639–1652.

Carstens B. C., and J. D. Satler. 2013. The carnivorous plant described as *Sarracenia alata* contains two cryptic species. Biol. J. Linn. Soc. 109:737–746.

Clayton, D. S., S. Bush, and K. Johnson. 2004. Ecology of congruence: past meets present. Syst. Biol. 53:165–173.

Dahlem G. A., and R. C. Naczi. 2006. Flesh flies (Diptera: Sarcophagidae) associated with North American pitcher plants (Sarraceniaceae), with descriptions of three new species. Ann. Ento. Soc. Am. 99:218–240.

Darwin, C. 1859. The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London, UK.

———. 1875. Insectivorous plants. John Murray, London, UK.

Dawson, M. N. 2013. Natural experiments and meta-analyses in comparative phylogeography. J. Biogeogr. 41:52–65.

Dellicour, S., and P. Mardulyn. 2014. SPADS 1.0: a toolbox to perform spatial analyses on DNA sequence data sets. Mol. Ecol. Resour. 14:647–651.

Dyer, R. J. 2012. The gstudio package. Virginia Commonwealth University, Virginia.

Edwards, S. V., and P. Beerli. 2000. Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution 54:1839–1854.

Edwards, S. V., L. Liu, and D. K. Pearl. 2007. High-resolution species trees without concatenation. Proc. Nat. Acad. Sci. 104:5936–5941.

Espindola, A., B. C. Carstens, and N. Alvarez. 2014. Comparative phylogeography of mutualists and the effect of the host on the genetic structure of its partners. Biol. J. Linn. Soc. 113:1021–1035.

Excoffier L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Farrell, B., and C. Mitter. 1990. Phylogenesis of insect/plant interactions: have *Phyllobrotica* leaf beatles (Chrysomelidae) and the Lamiales diversified in parallel? Evolution 44:1389–1403.

Felsenstein, J. 2006. Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? Mol. Biol. Evol. 23:691–700.

Foelix, R. F. 1982. Biology of spiders. Harvard Univ. Press, Harvard.

Folkerts, D. 1999. Pitcher plant wetlands of the southeastern United States. Pp. 247–275 *in* D. P. Batzer, R. B. Rader, and S. A. Wissinger, eds. Invertebrates in freshwater wetlands of North America: Ecology and management. John Wiley and Sons, Inc., New York, NY.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3:294–299.

Fouquet, A., B. P. Noonan, M. T. Rodrigues, N. Pech, A. Gilles, and N. J. Gemmell. 2012. Multiple quaternary refugia in the eastern guiana shield revealed by comparative phylogeography of 12 frog species. Syst. Biol. 61:461–489.

Garrick, R. C., D. M. Rowell, C. S. Simmons, D. M. Hillis, and P. Sunnucks. 2008. Fine-scale phylogeographic congruence despite demographic incongruence in two low-mobility sproxylic springtails. Evolution 62:1103–1118.

Greenstone, M. H., C. E. Morgan, A. Hultsch, R. A. Farrow, and J. E. Dowse. 1987. Ballooning spiders in Missouri, USA, and New South Wales, Australia: family and mass distributions. J. Arachnol. 15:163–170.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704.

Hafner, M. S., J. W. Demastes, T. A. Spradling, and D. L. Reed. 2003. Cophylogeny between pocket gophers and chewing lice. Pp. 195–218 *in* R. D. M. Page, ed. Tangled trees: Phylogeny, cospeciation, and coevolution. Chicago Univ. Press, Chicago.

Heled, J., and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27:570–580.

Herbert, P. D. N., S. Ratnasingham, and J. R. deWaard. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. Lond. B. 270:s96–s99.

Hickerson, M. J., E. Stahl, and N. Takebayashi. 2007. msBayes: pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. BMC Bioinform. 8:268.

Hickerson, M. J., B. C. Carstens, J. Cavender-Bares, K. A. Crandall, C. H. Graham, J. B. Johnson, L. Rissler, P. F. Victoriano, and A. D. Yoder. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Mol. Phylogenet. Evol. 54:291–301.

Hickerson, M. J., E. A. Stahl, and H. A. Lessios. 2006. Test for simultaneous divergence using approximate bayesian computation. Evolution 60:2435–2453.

Hope, A. G., S. Y. W. Ho, J. L. Malaney, J. A. Cook, and S. L. Talbot. 2014. Accounting for rate variation among lineages in comparative demographic analyses. Evolution 68:2689–2700.

Hudson, R. R. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics 18:337–338.

Hudson, R. R., and J. A. Coyne. 2002. Mathematical consequences of the genealogical species concept. Evolution 56:1557–1565.

Huang, W., N. Takebayashi, Y. Qi, and M. J. Hickerson. 2011. MTML-msBayes: approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. BMC Bioinform. 12:1.

Jensen, J. L., A. J. Bohonak, and S. T. Kelley. 2005 Isolation by distance, web service. BMC Genet. 6:13. v.3.23. Available at http://ibdws.sdsu.edu/

Johnson, M. T. J., and J. R. Stinchcombe. 2007. An emerging synthesis between community ecology and evolutionary biology. Trends Ecol. Evol. 22:250–257.

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

Larget, B. R., S. K. Kotha, C. N. Dewey, and C. Ané. 2010. BUCKy: Gene tree/species tree reconciliation with bayesian concordance analysis. Bioinformatics 26:2910–2911.

Jousselin, E., S. van Noort, V. Berry, J.-Y. Rasplus, N. Rønsted, J. C. Erasmus, and J. M. Greeff. 2008. One fig to bind them all: host conservatism in a fig wasp community unraveled by cospeciation analyses among pollinating and nonpollinating fig wasps. Evolution 62:1777–1797.

Knowles, L. L., and W. P. Maddison. 2002. Statistical phylogeography. Mol. Ecol. 11:2623–2635.

Koopman, M. M., and B. C. Carstens. 2010. Conservation genetic inferences in the carnivorous pitcher plant *Sarracenia alata* (Sarraceniaceae). Conserv. Genet. 11:2027–2038.

Kubatko, L. S., B. C. Carstens, and L. L. Knowles. 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25:971–973.

Maddison, D. R., and W. P. Maddison. 2005 MacClade 4: analysis of phylogeny and character evolution. Version 4.08a. Available at http://macclade.org.

McKenna, D. D., A. S. Sequeira, A. E. Marvaldi, and B. D. Farrell. 2009. Temporal lags and overlap in the diversification of weevils and flowering plants. Proc. Nat. Acad. Sci. 106:7083–7088.

McPherson, S. 2007. Pitcher plants of the Americas. The McDonald and Woodward Publishing Company, Blacksburgh, VA.

Muma, M. H., and Denmark, H. A. 1967. Biological studies on *Macroseius biscutatus* (Acarina: Phytoseiidae). Flor. Entom. 249–255. Q

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci. 70:3321–3323.

Nelson, G., and N. I. Platnick. 1981. Systematics and biogeography—cladistics and vicariance. Columbia Univ. Press, New York, New York.

Noss, R. F., E. T. LaRoe III, and J. M. Scott. 1995. Endangered ecosystems of the United States: a preliminary assessment of loss and degradation. Vol. 28. U. S. Department of the Interior, National Biological Service, Washington, DC, USA.

Oaks, J. R., J. Sukumaran, J. A. Esselstyn, C. W. Linkem, C. D. Siler, M. T. Holder, and R. M. Brown. 2012. Evidence for climate-driven diversification? A caution for interpreting ABC inferences of simultaneous historical events. Evolution 67:991–1010.

Oaks, J. R. 2014. An improved approximate-Bayesian model-choice method for estimating shared evolutionary history. BMC Evol. Biol. 14:150.

Oaks, J. R., C. W. Linkem, and J. Sukumaran. 2014. Implications of uniformly distributed, empirically informed priors for phylogeographical model selection: a reply to Hickerson et al. Evolution 68:3607–3617. Q

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2015. vegan: community ecology package. r package version 2.2–1. Available at http://CRAN.R-project.org/package=vegan

Page, T. J., and J. M. Hughes. 2014. Contrasting insights provided by single and multispecies data in a regional comparative phylogeographic study. Biol. J. Linn. Soc. 111:554–569.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Paradis, E. 2010. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics 26:419–420.

Pelletier, T. A., and B. C. Carstens. 2014. Model choice for phylogeographic inference using a large set of models. Mol. Ecol. 23:3028–3043.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25:1253–1256.

R Core Team. 2013. R: a language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria. Available at URL http://www.R-project.org/.

Rambaut, A., and N. C. Grass. 1997. Seq-gen: an application for the monte carlo simulation of DNA sequence evolution along phylogenetic trees. Comp. App. Biosci. 13: 235–238.

Rambaut, A., M. A. Suchard, D. Xie, and A. J. Drummond. 2014. Tracer v1.6, Available at http://beast.bio.ed.ac.uk/Tracer

Robert, C. P., J. Cornuet, J. Marin, and N. S. Pillai. 2011. Lack of confidence in approximate Bayesian computation model choice. Proc. Nat. Acad. Sci. 108:15112–15117.

Roe, A. D., A. V. Rice, D. W. Coltman, J. E. K. Cooke, and F. A. H. Sperling. 2010. Comparative phylogeography, genetic differentiation and con-

trasting reproductive modes in three fungal symbionts of a multipartite bark beetle symbiosis. Mol. Ecol. 20:584–600.

Rokas, A., and S. B. Carroll. 2005. More genes or more taxa? the relative contribution of gene number and taxon number to phylogenetic accuracy. Mol. Biol. Evol. 22:1337–1344.

Rønsted, N., G. D. Weiblen, J. M. Cook, N. Salamin, C. A. Machado, and V. Savolainen. 2005. 60 million years of co-divergence in the fig-wasp symbiosis. Proc. R. Soc. B 272:2593–2599.

Satler, J. D., A. J. Zellmer, and B. C. Carstens. 2016. Biogeographic barriers drive co-diversification within associated eukaryotes of the *Sarracenia alata* pitcher plant system. PeerJ 4:e1576.

Smith, C. I., S. Tank, W. Godsoe, J. Levenick, E. Strand, T. Esque, and O. Pellmyr. 2011. Comparative phylogeography of a coevolved community: concerted population expansions in Joshua Trees and four yucca moths. PLoS ONE 6:e25628.

Smith, B. T., J. E. McCormack, A. M. Cuervo, M. J. Hickerson, A. Aleixo, C. D. Cadena, J. Pérez-Emán, C. W. Burney, X. Xie, M. G. Harvey, et al. 2014a. The drivers of tropical speciation. Nature 515:406–409.

Smith, B. T., M. G. Harvey, B. C. Faircloth, T. C. Glenn, and R. T. Brumfield. 2014b. Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Syst. Biol. 63:83–95.

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

Stephens, J. D., S. R. Santos, and D. R. Folkerts. 2011. Genetic differentiation, structure, and a transition zone among populations of the pitcher plant moth *Exyra semicrocea*: implications for conservation. PLoS ONE 6:e22658.

Sukumaran, J., and M. T. Holder. 2010. DenroPy: a python library for phylogenetic computing. Bioinformatics 26:1569–1571.

Sukumaran, J., and M. T. Holder. 2015. SumTrees: phylogenetic tree summarization. 4.0.0. Available at https://github.com/jeetsukumaran/DendroPy.

Swenson, N. G., and D. J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. Am. Nat. 166:581–591.

Wegmann, D., C. Leuenberger, S. Neuenschwander, and L. Excoffier. 2010. ABCtoolbox: a versatile toolkit for approximate Bayesian computations. BMC Bioinform. 11:116.

Whiteman, N. K., R. T. Kimball, and P. G. Parker. 2007. Co-phylogeography and comparative population genetics of the threatened Galápagos hawk and three ectoparasite species: ecology shapes population histories within parasite communities. Mol. Ecol. 16:4759–4773.

Zellmer, A. J., M. M. Hanes, S. M. Hird, and B. C. Carstens. 2012. Deep phylogeographic structure and environmental differentiation in the carnivorous plant *Sarracenia alata*. Syst. Biol. 61:763–777.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

Associate Editor: F. Burbrink
Handling Editor: M. Servedio

Supporting Information

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

Figure S1. Spatial principle components analysis (sPCA) of the six arthropods.
Figure S2. Maximum likelihood COI mtDNA gene tree for *Sarcophaga sarracenia*.
Figure S3. Maximum likelihood COI mtDNA gene tree for *Fletcherimyia celarata*.
Figure S4. Maximum likelihood COI mtDNA gene tree for *Exyra semicrocea*.
Figure S5. Maximum likelihood COI mtDNA gene tree for *Macroseius biscutatus*.
Figure S6. Maximum likelihood COI mtDNA gene tree for *Peucetia viridans*.
Figure S7. Maximum likelihood COI mtDNA gene tree for *Misumenoides formosipes*.
Figure S8. Distribution of average PCF values when pairwise comparisons are made between species simulated from the same tree.
Figure S9. Maximum likelihood gene tree (*rps*16–*trn*K) for *Sarracenia alata*.
Table S1. Arthropod sampling information for COI mtDNA and corresponding GenBank accession numbers for each individual.
Table S2. RDA analysis. Results suggests that geography explains genetic variation in all species minus the spiders (based on significant p-values).
Table S3. Phylogeographic concordance factors for all permutations of K levels from the Pacific Northwest data set.
Table S4. Phylogeographic concordance factors for all permutations of K levels from the Neotropical bird data set.
Table S5. Results from PyMsBayes analyses.
Table S6. Results from PyMsBayes analyses with spiders removed.