American Journal of Botany

RESEARCH ARTICLE

Phylogeography of a widespread eastern North American shrub, *Viburnum lantanoides*

Brian Park[1,2,3] and Michael J. Donoghue[1]

Manuscript received 5 July 2018; revision accepted 10 December 2018.

[1] Department of Ecology and Evolutionary Biology, Yale University, P.O. Box 208106, New Haven, Connecticut 06520, USA

[2] Current address: Department of Plant Biology, University of Georgia, Miller Plant Sciences Building, Athens, Georgia 30602, USA

[3] Author for correspondence (e-mail: brian.park@uga.edu)

Citation: Park, B. and M. J. Donoghue. 2019. Phylogeography of a widespread eastern North American shrub, *Viburnum lantanoides*. *American Journal of Botany* 106(3): 389–401.

doi:10.1002/ajb2.1248

PREMISE OF THE STUDY: There have been relatively few phylogeographic studies of eastern North American plants, especially of animal-dispersed shrubby species, and this leaves a significant gap in our understanding of how such species were affected by glacial events. Here, we analyzed the phylogeography of the widespread understory shrub *Viburnum lantanoides*.

METHODS: We generated RADseq data and paleoclimatic species distribution models (SDMs) to identify the locations of refugia where *V. lantanoides* may have survived the Last Glacial Maximum (LGM) and how its range expanded as glaciers receded.

KEY RESULTS: Genetic diversity falls off with increasing latitude and longitude, indicating that range expansion likely occurred via serial founder events from southern source populations. Samples from the southern Appalachians form a grade, while those from the north form a clade, suggesting that a single genetic lineage recolonized the north. SDMs indicate that *V. lantanoides* probably survived the LGM in refugia on the mid-Atlantic Coastal Plain and/or the interior Gulf Coastal Plain.

CONCLUSIONS: Our analyses indicate that *V. lantanoides* survived the LGM in refugia south of the glacier but north of the extensive refugium along the Gulf Coast. Following the LGM, a single population expanded northward along the Appalachian Mountains and eventually into eastern Canada. The patterns observed here suggest that range expansion occurred in a stepwise manner, similar to postglacial dynamics observed in a number of European plant species.

KEY WORDS eastern North America; Last Glacial Maximum; RADseq; species distribution models.

Repeated episodes of glacial expansion and retreat, most recently during the Last Glacial Maximum (LGM; ~21,500 ybp), have had a significant impact on the evolutionary histories of plant species in the temperate Northern Hemisphere (Hewitt, 2000, 2004). Regional phylogeographic studies have been instrumental in identifying similarities and differences in how species responded to these events. These have provided important insights into the locations of refugia where species persisted during glacial periods and the modes of postglacial range expansion that explain the distribution of genetic diversity over their contemporary geographic ranges.

A number of important phylogeographic patterns have emerged from these studies. In Europe, where climatic conditions were especially harsh during the LGM, many temperate forest species exhibit declines in genetic diversity at higher latitudes, which has been explained by northward range expansions from southern refugia characterized by long-distance dispersal and founder events (Hewitt, 1999, 2000). Some European species have been found to be most genetically diverse at midlatitudes, which may be a consequence of secondary contact between isolated populations originating from separate southern refugia (Petit, 2003). In Eastern Asia, by contrast, climatic conditions during the LGM were less severe, and many temperate forest species may have survived the LGM more or less in place, perhaps in multiple refugia distributed throughout their contemporary ranges (e.g., Hu et al., 2008; Tian et al., 2009; Bai et al., 2010; Zeng et al., 2015). In western North America, the complex physiography of the region combined with the presence of large refugia in Beringia and the Pacific Northwest, as well as small refugia on coastal islands, has generated north-south, south-north, and east-west patterns of genetic diversity and population structure (Brunsfeld et al., 2001; Shafer et al., 2010).

Phylogeographic patterns in temperate forest species of eastern North America (ENA) are not as well defined, which may reflect the fact that large portions of unglaciated ENA were habitable during the LGM (Jackson et al., 2000; Williams et al., 2004). Some temperate forest species that currently span glaciated and unglaciated ENA conform to the European model of rapid, northward range expansion from southern refugia (e.g., Schwaegerle and Schaal, 1979; Lewis and Crawford, 1995; Broyles, 1998; Magni et al., 2005). However, other species with similarly wide ranges have been found to lack such spatial patterns in genetic diversity, suggesting that postglacial range expansion occurred gradually along an expansive migration front (e.g., Lumibao et al., 2017; Bemmels and Dick, 2018). To complicate matters, some species are hypothesized to have survived the LGM in ice-free, coastal areas well above the southern extent of the ice sheets (e.g., Godbout et al., 2010; Suarez-Gonzalez et al., 2015). Taken together, these findings demonstrate that the biogeographic history of ENA has been complicated, and that the recovery of a common phylogeographic pattern for the postglacial history of ENA forest species is unlikely.

In considering these patterns, it is important to note that most phylogeographic studies of ENA plants have been conducted in wide-ranging, wind-pollinated trees, with fruits/seeds that are not dispersed by birds: for example, *Fagus grandifolia* (Morris et al., 2010), *Liquidambar styraciflua* (Morris et al., 2008), *Pinus strobus* (Nadeau et al., 2015; Zinck and Rajora, 2016), and species in the genera *Acer* (McLachlan et al., 2005; Saeki et al., 2011; Vargas-Rodriguez et al., 2015), *Betula* (Thomson et al., 2014), and *Carya* (Bemmels and Dick, 2018). Despite their diverse ecologies, these plants are long lived, with generally large population sizes (Savolainen et al., 2007; Krutovsky et al., 2012; Alberto et al., 2013). By contrast, only a handful of studies have been conducted on ENA herbs—for example, *Asclepias exaltata* (Broyles, 1998), *Polygonella* spp. (Lewis and Crawford, 1995), *Sarracenia purpurea* (Schwaegerle and Schaal, 1979), *Solidago* subsect. *Humiles* (Peirson et al., 2013), and *Trillium* spp. (Griffin and Barrett, 2004a, b; Gonzales et al., 2008)—and very few have been conducted on understory shrubs (e.g., *Dirca palustris* [Peterson and Graves, 2015] and the *Viburnum nudum* and *V. lentago* species complexes [Spriggs et al., 2019, unpublished data]). Many of these plants are patchily distributed and rely on animals for both pollination and dispersal, and they are therefore likely to have responded differently to glacial episodes than the trees that have been studied to date.

Given this gap in knowledge, detailed studies of additional widespread shrubby species could provide new insights into how plants responded to the LGM. To this end, here we combine the use of reduced-representation sequencing methods (RADseq) with species distribution modeling to reconstruct the phylogeographic history of the hobblebush (*Viburnum lantanoides* Michx., Adoxaceae). The goals of our study were to assess patterns of genetic diversity across the wide geographic range of this species, to identify the location of refugial areas where it may have survived the LGM, and to infer how it might have expanded northward following glaciation.

MATERIALS AND METHODS

Study species

Viburnum lantanoides is an understory shrub that inhabits cool, mesic habitats in mixed hardwood and boreal forests from the southern Appalachians to the Canadian Maritime Provinces. In the northern portion of its range, *V. lantanoides* can be locally abundant, often occurring in dense thickets in cool moist forests throughout northern New England, the Canadian Maritimes to the east, and southeastern Ontario to the west. It is more patchily distributed in the southern portion of its range, occurring exclusively at higher elevations in spruce–fir forests in the southern Appalachians. Flowers are borne in compound umbel-like inflorescences in which a cluster of small perfect flowers is encircled by sterile flowers with greatly enlarged corollas. Plants flower for a short time during early spring and are visited primarily by andrenid bees, syrphid flies, and elaterid beetles (Park et al., 2019). Fruits are mainly produced through outcrossing (Park et al., 2019) and typically are dispersed by birds, although rates of frugivory are low (Gould, 1966).

Sampling and sequencing

One to three plants were sampled from 69 localities distributed throughout the geographic range of *V. lantanoides* (Appendix S1). Latitude/longitude coordinates for each locality were obtained using a Garmin GPSmap 62s handheld GPS unit (Olathe, Kansas, USA) to 2–2000 m precision. Leaves were preserved in silica gel, and voucher specimens for each population were deposited in the Yale University Herbarium (YU). A modified CTAB method was used to extract genomic DNA from our collections (Healey et al., 2014). The resulting DNA extractions were prepared for RAD-sequencing (Baird et al., 2008) with the restriction enzyme *Pst1* by Floragenex (Portland, Oregon, USA). RAD-libraries (two libraries consisting of 95 samples each) were sequenced at the GC3F sequencing facility at the University of Oregon (Eugene, Oregon, USA) over two lanes of Illumina Hi-Seq 4000 to produce 100 bp, single-end reads.

Sequence processing

The Python program iPyrad version 0.7.12 (Eaton, 2014) was used to demultiplex, filter, and generate homologous de novo RAD-loci. Raw reads were demultiplexed to individual samples by unique 10 bp sample-specific barcodes and filtered to remove adapter-contaminated reads and reads with more than four low-quality bases (phred scores <20). Filtered reads were clustered within samples at 88% sequence similarity into de novo loci, and loci with fewer than six or >10,000 reads were removed from further processing. Resulting loci were clustered across samples at 88% sequence similarity and processed into a "base" assembly, where all loci were shared between four or more ingroup and outgroup samples. In total, we included 190 *V. lantanoides* accessions. For rooting purposes we included one sample each of two closely related species within the *Pseudotinus* clade (*V. sympodiale* of China and Taiwan, and *V. furcatum* of Japan). Twelve samples with <10,000 loci were excluded from further processing because they were found to contribute disproportionately to the amount of missing data in downstream analyses.

The "base" assembly was then processed to generate two assemblies for analysis: a "phylogenomics" assembly for phylogenetic analyses, where all loci were shared among ≥90 ingroup and outgroup samples (178 *V. lantanoides* samples, two outgroup samples); and an "ingroup" assembly for population genomic analyses, where all loci were shared among ≥89 ingroup samples (n = 178 samples). The "ingroup" assembly was subsequently filtered using the R package "radiator" (Gosselin, 2017) to randomly select one single-nucleotide polymorphism (SNP) per RAD-locus, to remove

SNPs with more than two alleles (*V. lantanoides* is diploid, with $2N = 18$; Egolf, 1962), and to remove SNPs with minor allele frequency <0.01. Our relatively low sample coverage threshold was chosen because it allowed for the inclusion of high-mutation-rate loci, which may be useful in resolving relationships among more recently diverged populations (Huang and Knowles, 2014; Eaton et al., 2017), while allowing for reasonable levels of missing data, an excess of which may bias phylogeographic signal (Rubin et al., 2012).

Genetic diversity

Global population genetic parameters (e.g., F_{is}, H_o, H_t, H_s) were estimated using the R package "hierfstat" (Goudet, 2005). Overall F_{st} was estimated using Weir and Cockerham's (1984) method in the R package "assigner" (Gosselin et al., 2016), with 100 bootstrap replicates to assess significance. We assessed patterns of isolation-by-distance (IBD) by performing a Mantel test using the "mantel.randtest" function from the R package "ade4" (Dray and Dufour, 2007) and 1×10^6 Monte Carlo simulations to test for significance. We calculated pairwise genetic distances between samples (i.e., the proportion of allelic differences) using the "bitwise.dist" function in the R package "poppR" (Kamvar et al., 2014), and geographic distances between collection localities using the "distm" function in the R package "geosphere" (Hijmans, 2016). We regressed sample observed heterozygosity (H_o) against latitude and longitude to identify geographic patterns in the distribution of genetic diversity. We used BCFtools (Li et al., 2009) to estimate H_o by dividing the number of heterozygous sites by the sum of homozygous and heterozygous sites sequenced per sample. Finally, we performed a hierarchical analysis of molecular variance (AMOVA) in "poppR," with 10,000 permutations to assess significance. AMOVA was performed for the following levels: between genetic clusters (described below); among localities within genetic clusters; among samples at each collecting locality; and within samples.

Population structure

Four approaches were used to assess population structure. We first used sNMF (Frichot et al., 2014) as implemented in the R package "LEA" (Frichot and François, 2015) to assign samples to genetic clusters. sNMF estimates individual ancestry coefficients utilizing the same likelihood model underlying Structure (Pritchard et al., 2000) and Admixture (Alexander et al., 2009), but uses non-negative matrix factorization and least-squares optimization to accommodate large, genome-scale datasets. sNMF is also robust to deviations from Hardy-Weinberg equilibrium, making it an effective tool to infer population structure in organisms with diffuse population structure. We complemented the sNMF analyses with tess3r (Caye et al., 2016), which utilizes the same algorithm as sNMF but incorporates geographic coordinates in estimating sample ancestry coefficients. Both sNMF and tess3r analyses were performed using default parameters for $K = 1–10$, with 10 repetitions for each value of K, and the optimal value of K assessed using the cross-entropy criterion. We also used two model-free approaches to infer genetic clusters. We used discriminant analysis of principal components (DAPC; Jombart et al., 2010) as implemented in the R package "adegenet" (Jombart, 2008), using K-means clustering to identify the optimal number of clusters from $K = 1–10$, and estimating individual admixture coefficients using $n/3$ principal components ($n = 178$ samples). We then calculated Bayesian information criterion scores for each value of K to identify the best-fit number of clusters. We also conducted principal component analysis (PCA) as implemented in the R package "LEA" to visualize samples in two-dimensional genetic space.

Phylogenetic relationships

Supermatrix and species-tree approaches were used to infer phylogenetic relationships among samples in the "phylogenomic" assembly. Supermatrix analyses were conducted using RAxML version 8.2.1 (Stamatakis, 2014) under the GTR + CAT substitution model with 100 rapid bootstrap replicates to assess support. Species-tree analyses were conducted in "tetrad," an implementation of the program SVDquartets (Chifman and Kubatko, 2014) in iPyrad; tetrad estimates topologies for a user-specified number of quartets present in a set of taxa and assembles the resulting quartets into a supertree. For our analyses, we sampled all quartets and assessed node support with 100 nonparametric bootstrap replicates.

Demographic modeling

We used demographic model testing to identify the scenario that best describes the demographic history of *V. lantanoides*. Based on the results of our population structure and phylogenetic analyses, we focused on determining the timing of the origin of the northern lineage and its mode of population growth. Our models were based on a simple scenario in which ancestral lineages of *V. lantanoides* differentiate into two independently evolving entities: a southern Appalachian cluster comprising several lineages from the southern Appalachian mountains and a northern lineage representing the single genetic lineage that appears to have colonized the northern portion of the range of *V. lantanoides* (see below). Models were then formulated to identify the mode of postdivergence population growth, where we allowed for nonlinear changes in effective population size in none (Appendix S2A), one (Appendix S2B, S2C), or both (Appendix S2D) the southern Appalachian cluster and the northern lineage. Our models estimated three parameters: contemporary effective population sizes (Ne) of the southern Appalachian cluster (N1) and northern lineage (N2), and divergence time for the northern lineage (T). We did not consider island models (Wright, 1943) because the significant pattern of IBD and diffuse population structure are poor fits to *V. lantanoides* demographic history. Furthermore, considering the monophyly of the northern samples and a significant pattern of IBD, we did not include a migration parameter in our models.

We fit models using the diffusion approximation approach in δaδi (Gutenkunst et al., 2009). We generated a folded, two-dimensional joint site frequency spectrum from the "ingroup" assembly using the Python script "easySFS" (https://github.com/isaacovercast/easySFS), sampling 14 alleles from the southern Appalachian cluster and 16 alleles from the northern lineage to maximize the number of segregating sites for the analysis. Initial optimizations were conducted using 20 independent replicates of twofold perturbed parameters and were optimized using the L-BFGS-B algorithm for a maximum of 100 repetitions. Parameter estimates from the replicates with the highest log-likelihood were used to generate 20 additional sets of onefold perturbed starting points for final optimization, and the repetition with the highest log-likelihood was used to compare models and estimate parameter uncertainties.

We compared models using the Akaike information criterion (AIC). Calculation of AIC scores and model selection was conducted following Burnham and Anderson (2007). Standard deviations for the maximum-likelihood parameter estimates were obtained using the Fisher information matrix (FIM) approach (Coffman et al., 2016), which is a computationally efficient alternative to traditional bootstrapping. The FIM approach may not be appropriate for datasets composed of highly linked SNPs or when fitting complex models. However, considering that we sampled one SNP per RAD-locus, and that our models estimated four free parameters, we believe that the FIM approach provided reasonable uncertainty estimates for our analysis. Parameter estimates and uncertainties were transformed into biological units from model-derived estimates of Θ (population mutation rate). Mutation rates in *Viburnum* are unknown, so we conservatively assumed a rate obtained for the woody plant *Populus* (2.5×10^{-9} substitutions site^{-1} yr^{-1}; Ingvarsson, 2008). We also assumed a generation time of 10 yr, which we based on the observation that *V. lantanoides* produces "double-dormant" seeds (i.e., seeds need to be exposed to two successive winters before emergence of the epicotyl) and that several years of growth are needed before plants become reproductive.

Species distribution models

We used MaxEnt version 3.3.3k (Phillips et al., 2004, 2006) as implemented in the R package "dismo" (Hijmans et al., 2017) to estimate the current and historical geographic range of *V. lantanoides*. Data for 19 bioclimatic variables under current climates (1950–1990) and three LGM climate models (22 kya; CCSM-4, MIROC, and MPI-ESM) were downloaded at 2.5 arcminute resolution (~5 km^2 at the equator) from WorldClim (https://www.worldclim.org). The three LGM models are based on different general circulation models that vary in how they simulate climate dynamics (Varela et al., 2015). We pruned the climate data used to estimate distribution models by sampling 10,000 random points spanning our study region (latitude: 25–58°N, longitude: 100–50°W) and removing variables with a Pearson $R^2 > 0.7$. Using this approach, we retained the following six variables: mean annual temperature (bio1), mean diurnal range (bio2), mean temperature of the wettest quarter (bio8), mean annual precipitation (bio12), precipitation seasonality (bio15), and precipitation of the warmest quarter (bio18).

Locality data were from our own collections, which cover much of the known range of *V. lantanoides* (Appendix S3; $n = 86$ localities). We chose not to include data from other sources (e.g., Global Biodiversity Information Facility) because those data are less reliable and contain misidentifications. We sampled 10,000 pseudo-absence points (Barbet-Massin et al., 2012) from a minimum-convex polygon drawn two decimal degrees around each occurrence. Distribution models were generated using current climate data and then evaluated using fivefold cross-validation and the area-under-curve (AUC) statistic. Finally, a current climate model using all occurrences was used to project the distribution of *V. lantanoides* using the three models of LGM climate.

RESULTS

RADseq data

We generated 5.45×10^8 single-end, 100 bp reads, of which 4.53×10^8 reads (83%) were assigned to individual samples. On average, 2.39×10^6 reads were sequenced per sample (range = $2.34 \times 10^5 - 8.07 \times 10^6$ reads). After quality filtering, the "phylogenomics" assembly was composed of 35,796 RAD-loci and 44,743 parsimony-informative sites, each sequenced in ≥90 samples (of 178 *V. lantanoides*, one *V. sympodiale*, and one *V. furcatum*). The "ingroup" assembly was composed of 11,648 SNPs sampled from the same number of RAD-loci, each sequenced in a minimum of 89 samples (of 178 *V. lantanoides*). The total amounts of missing data were 39.8% and 41.8% in the "phylogenomic" and "ingroup" assemblies, respectively.

Genetic diversity

We detected moderate population differentiation ($F_{st} = 0.186$, 95% confidence interval [CI] = 0.178–0.195; Appendix S4) and high rates of inbreeding ($F_{is} = 0.487$; Appendix S4) among populations of *V. lantanoides*. A Mantel test revealed a significant negative relationship between pairwise geographic and genetic distance (Fig. 1A: $R^2 = 0.274$, $P < 0.0001$), indicating a significant pattern of IBD. There was a significant decline in observed heterozygosity (H_o) with increasing latitude (Fig. 1B: $R^2 = 0.377$, $F_{1,176} = 108$, $P < 0.0001$) and decreasing longitude (Fig. 1C: $R^2 = 0.245$, $F_{1,176} = 58.5$, $P < 0.0001$). AMOVA indicates that only a small proportion of genetic diversity was explained by differences between clusters (6.38%; Appendix S5) and localities within clusters (14.7%; Appendix S5). Instead, the majority of variation was found between samples within localities (36%; Appendix S5) and within samples (42.9%; Appendix S5). The observation of moderate range-wide genetic structure, significant IBD, and the partitioning of most genetic diversity at finer scales suggests that genetic diversity is continuously distributed across the geographic range of *V. lantanoides*, and that diversity decreases in a predictable manner with increasing distance away from the southern edge of its range.

Population structure

Both sNMF and DAPC analyses favored $K = 2$ genetic clusters. Using tess3r, cross-entropy scores decreased with increasing values of K and identified the optimal K as the number of sampled populations ($n = 69$), reflecting the significant pattern of IBD. As such, it is likely that $K = 2$ best describes the underlying genetic structure within *V. lantanoides*. The two clusters roughly correspond to a southern Appalachian cluster consisting of samples from North Carolina, Tennessee, Virginia, and West Virginia, and a northern cluster comprising samples from north of this region.

Interestingly, sNMF and tess3r identified diffuse population structure, with few samples possessing 100% membership in either cluster, except on the extreme edges of the range of *V. lantanoides* (e.g., North Carolina and eastern Nova Scotia/northeastern New Brunswick; Fig. 2B). sNMF, tess3r, and DAPC all identified "admixed" samples in Pennsylvania, suggesting that this area demarcates the boundary between the two lineages. Targeted analyses within the two clusters using all three methods did not identify significant substructure. The PCA supported diffuse population structure, with the position of samples in the PCA space being a function of their geographic locations (i.e., samples from nearby localities cluster together) with no discrete breaks between clusters (Fig. 2C, D). These results, taken together, support a pattern of IBD.

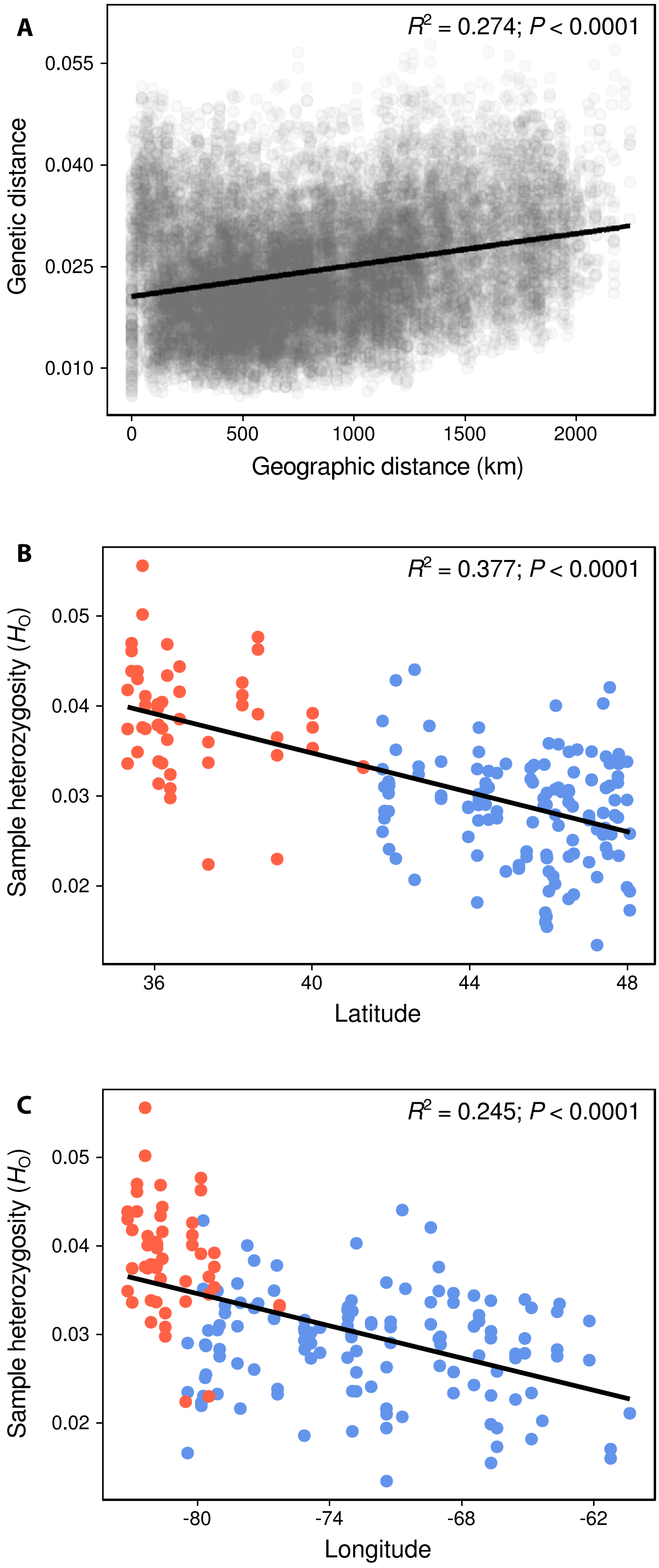

FIGURE 1. Spatial patterns in the distribution of genetic diversity in *Viburnum lantanoides*. Pairwise genetic distance increases with distance between collecting localities (A); sample observed heterozygosity decreases with increasing latitude (B) and also with longitude (C). Red circles denote samples with >50% southern Appalachian ancestry as inferred through population structure analyses. Similarly, blue circles denote samples of northern ancestry.

Phylogenetic relationships

Both concatenated and species-tree phylogenies suggest a south-north pattern of lineage splitting, where samples collected from the southern Appalachian region (North Carolina, Tennessee, Virginia) form a grade in relation to a clade composed of northern samples from West Virginia, Pennsylvania, New York, New England, Ontario, Quebec, Nova Scotia, and New Brunswick (Fig. 3). Concatenated and species-tree analyses yielded identical topologies, with only slight differences in node support values. The monophyly of the northern samples is consistent with the hypothesis that a single genetic lineage was responsible for recolonizing the northern portion of the *V. lantanoides* range following glacial retreat.

Demographic modeling

Across all models, effective population sizes in the southern Appalachian cluster and the northern lineage were estimated to be markedly smaller than the ancestral lineages (Nref; Table 1), suggesting a reduction in effective population size (i.e., genetic diversity) following the establishment of the northern lineage. Furthermore, the northern lineage was found to have originated between 38.2 and 45.6 kya (Table 1), suggesting that the differentiation between the southern Appalachian cluster and the northern lineage occurred prior to the LGM. Model selection using AIC favored a model with changes in effective population size in both the southern Appalachian cluster and the northern lineage (Appendix S2D), suggesting that reductions in effective population sizes may have occurred following the establishment of the latter. The magnitude of this change in effective population size was most apparent in this model, where there was a nearly twofold reduction between extant and ancestral lineages (Table 1). According to this model, the southern Appalachian cluster and the northern lineage are estimated to have diverged 45.6 kya (95% CI: 41.8–49.3 kya), nearly 20 kya before the LGM.

Distribution modeling

Distribution models had high predictive ability (AUC = 0.855 ± 0.005 [mean ± SD]) and accurately identified the northern and southern edges of the current distribution of *V. lantanoides* (Fig. 4A). The lack of large, contiguous areas with very high climate suitability accurately reflects the patchy distribution of *V. lantanoides* throughout its range. However, the model also overestimated the longitudinal range of *V. lantanoides*; suitable climates were identified along the northern shore of Lake Superior and in Newfoundland, areas where *V. lantanoides* has never been collected.

LGM projections varied considerably depending on the climate reconstruction used (Fig. 4B–D). Using CCSM4, *V. lantanoides* is predicted to have survived the LGM at midlatitudes on the exposed coastal shelf off North Carolina in the Atlantic Coastal Plain (Fig. 4B). By contrast, with MIROC, suitable climates were identified at lower latitudes over a larger area centered around

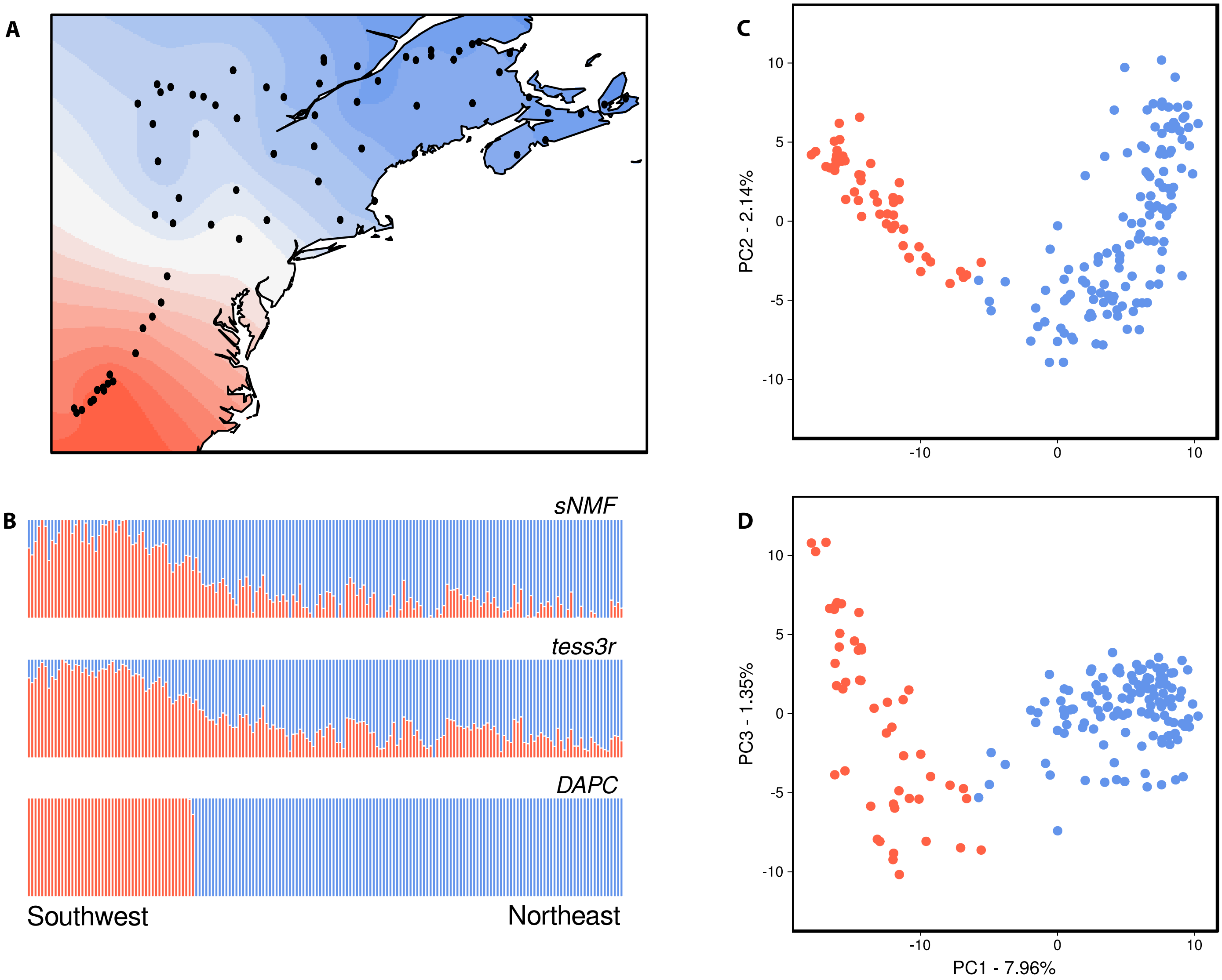

FIGURE 2. Population structure of *Viburnum lantanoides*. (A) Map depicting the distribution of ancestry coefficients inferred through tess3r for $K = 2$ clusters. Sampling localities are represented by black dots. More saturated colors indicate a greater proportion of ancestry to either cluster. (B) Individual admixture coefficients at $K = 2$ clusters inferred using sNMF, tess3r, and DAPC. PCA plots of the first and second (C) and of the first and third (D) principal components show that samples cluster by geography and that there is not a clear gap between southern Appalachian and northern populations.

current-day southern Arkansas, Mississippi, northern Louisiana, and eastern Texas in the interior Gulf Coastal Plain (Fig. 4C). With MPI, suitable climates were identified in both the Atlantic Coastal Plain and in the Gulf Coastal Plain in Arkansas and northern Louisiana (Fig. 4D). Despite these differences, all models indicate that *V. lantanoides* survived the LGM in a southern refugium, below the southern extent of the Laurentide Ice Sheet.

DISCUSSION

We reconstructed the postglacial migration history of *V. lantanoides* using recently developed genomic tools and paleoclimatic distribution modeling. We identified a significant pattern of IBD (Fig. 1A) and found that genetic diversity declined steadily with increasing distance from the southern portion of the geographic range (Fig. 1B, C). Samples sorted into one genetic cluster including the southern Appalachian populations and another one including the northern populations (Fig. 2A–D). However, overall population structure was diffuse and the proportion of southern Appalachian ancestry decreased steadily with increasing latitude and decreasing longitude. These results strongly support a pattern of IBD. Samples from the southern Appalachian region were found to form a grade while samples from north of this region formed a well-supported clade (Fig. 3), suggesting that a single genetic lineage may have colonized the northern portion of the range of *V. lantanoides*. The genetic differentiation of the southern Appalachian lineages from the northern lineage was inferred to have occurred prior to the LGM, with all having undergone significant reductions in effective population size following

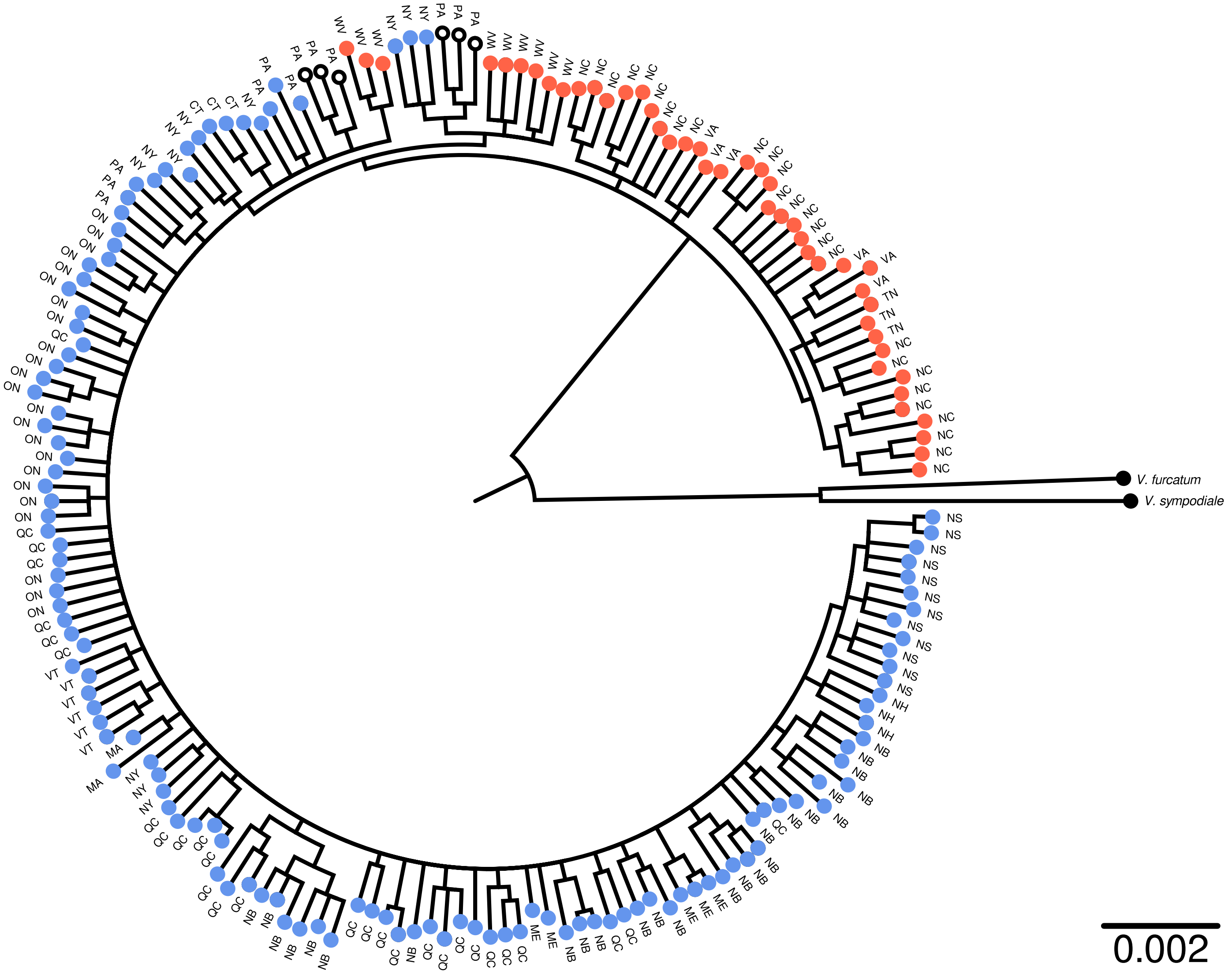

FIGURE 3. Phylogram depicting relationships among samples inferred through RAxML suggests that a single genetic lineage recolonized the northern portion of the range of *Viburnum lantanoides*. Red tips denote samples with a high proportion of southern Appalachian ancestry as inferred through population structure analyses; blue tips denote samples with a high proportion of northern ancestry; white tips denote "admixed" samples with nearly equal proportions of southern Appalachian and northern ancestry. Tip labels refer to the state or province the sample was collected in. All nodes were resolved with >90 bootstrap support.

the origination of the northern lineage (Table 1). Finally, using paleoclimate distribution modeling, we found that *V. lantanoides* underwent major range displacement during the LGM, and likely survived in refugial areas on the Atlantic Coastal Plain or the interior Gulf Coastal Plain or both (Fig. 4). Taken together, these analyses imply that *V. lantanoides* survived glacial periods in one or more southern refugia and then expanded its range into the southern Appalachian Mountains and then northward and ultimately eastward, in a series of founder events accompanied by the loss of genetic diversity.

Postglacial range expansion

Significant IBD (Fig. 1A), moderate population differentiation (F_{st} = 0.186), and an absence of strong geographic breaks in population structure (Fig. 2A–D) suggest that genetic variation is continuously distributed across the range of *V. lantanoides*. However, genetic variation was found to decrease predictably with increasing distance from the southern edge of its range (Fig. 1B, C), indicating that recolonization occurred in a northward and then eastward direction from source populations in the southern Appalachians. These patterns are suggestive of *V. lantanoides* having undergone a "stepping-stone" mode of postglacial range expansion (Sage and Wolff, 1986; Broyles, 1998), where recolonization of previously glaciated areas occurred through serial founder events along a single migration front, each of which drew propagules from an increasingly limited pool of genetic diversity.

Several lines of evidence support this scenario. Phylogenetic resolution of a southern Appalachian grade and a well-supported northern clade indicates that a single lineage dispersed out of the

TABLE 1. Maximum-likelihood estimates and 95% confidence intervals (CI) of population sizes and divergence times for *Viburnum lantanoides* inferred through demographic reconstruction. N1 refers to the southern Appalachian cluster; N2 refers to the northern lineage.

Model	AIC	ΔAIC	Log-likelihood	Nref (10^3) (95% CI)	N1 (10^3) (95% CI)	N2 (10^3) (95% CI)	T (10^3 yr) (95% CI)
(A) No size change	818	20.0	−405	26.1 (24.6–27.6)	12.7 (9.76–15.7)	6.13 (4.70–7.55)	38.2 (33.7–42.7)
(B) Size change in N1	810	12.3	−401	26.1 (24.7–27.5)	7.86 (5.20–10.5)	6.27 (5.86–6.68)	39.3 (35.1–43.5)
(C) Size change in N2	806	8.43	−399	25.7 (24.4–27.0)	14.1 (11.6–16.6)	3.14 (2.73–3.55)	44.5 (40.8–48.2)
(D) Size change in N1 and N2	798	0	−395	25.7 (24.4–27.0)	9.33 (6.72–11.9)	3.23 (2.95–3.52)	45.6 (41.8–49.3)

southern portion of the range (Fig. 3). Furthermore, high support for a demographic model with population expansion/contraction in the northern lineage (Appendix S2D; Table 1) and the threefold difference in effective population size estimated between the southern Appalachian cluster and the northern lineage (~9000 vs. ~3000; Table 1) signal that range expansion was accompanied by the loss of genetic diversity.

Overall, *V. lantanoides* follows the classic phylogeographic pattern of "southern diversity, northern purity" observed in many European temperate forest species (Hewitt, 1999, 2000, 2004; Petit, 2003). This pattern has also been demonstrated in a number of ENA species (e.g., Schwaegerle and Schaal, 1979; Broyles, 1998; Zinck and Rajora, 2016), though some ENA species appear to lack a clear clinal pattern in genetic diversity (e.g., Lumibao et al., 2017; Bemmels and Dick, 2018). This may reflect the relatively simple landscape and mild glacial history of ENA as compared with Europe (Hewitt, 2000; Shafer et al., 2010; Qiu et al., 2011), which may have enabled some species—especially wide-ranging, wind-pollinated trees—to have survived the LGM in large numbers over wide areas, and to have expanded their ranges gradually along a continuous migration front. This mode of postglacial range expansion would serve to preserve genetic diversity and maintain connectivity

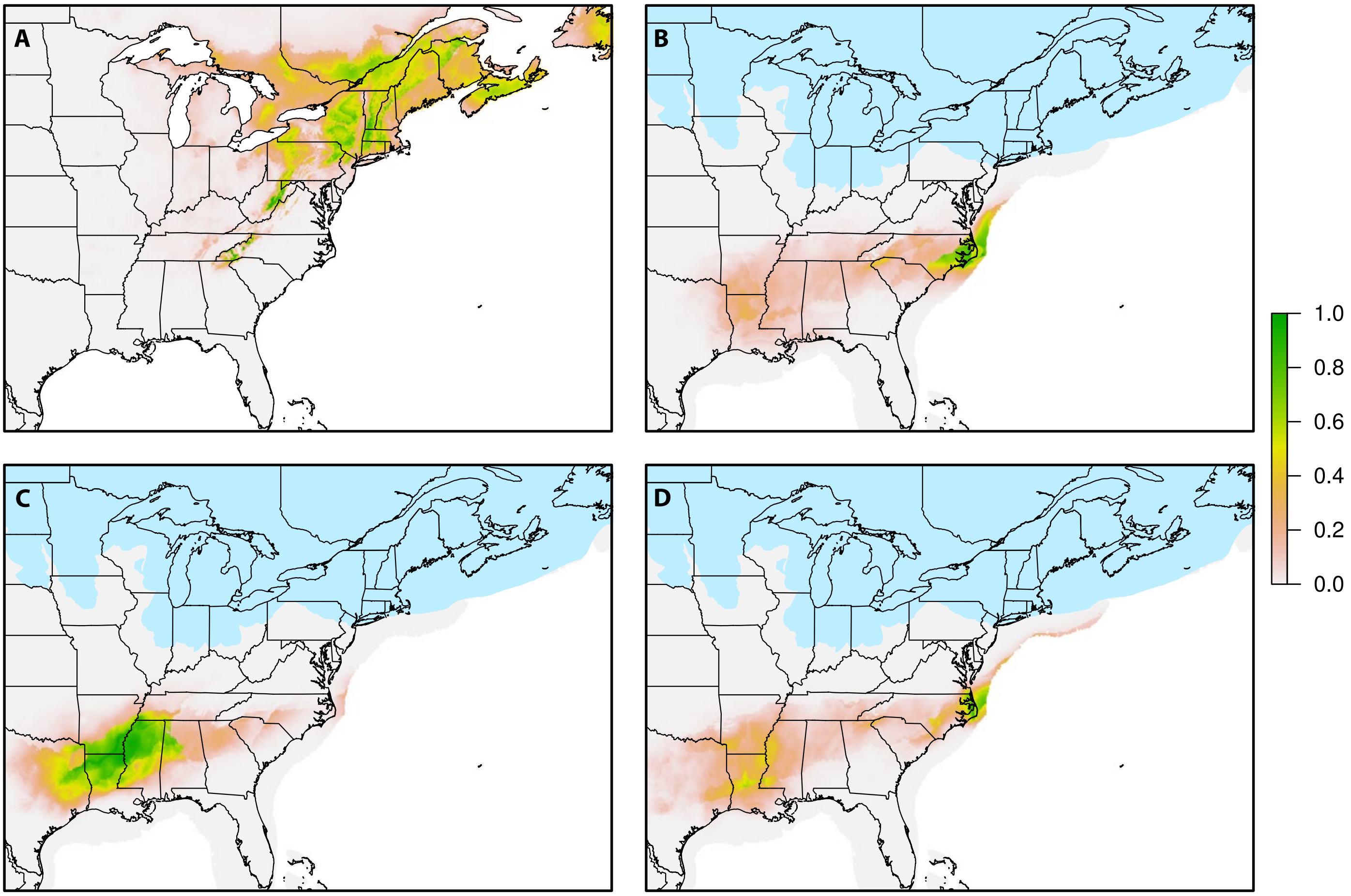

FIGURE 4. Species distribution models for *Viburnum lantanoides* inferred using Maxent. (A) Current climatic conditions. (B–D) Projections into LGM climate models: (B) CCSM4, (C) MIROC-ESM, and (D) MPI-ESM-P. Blue areas denote ice-covered areas during the LGM.

among populations (Ibrahim et al., 1996; Jaramillo-Correa et al., 2009). However, this is unlikely to be representative of patchily distributed understory shrubs, many of which occur in small and more isolated populations and have a more limited capacity for long-distance gene flow. This is evident in *V. lantanoides*, where most genetic variation (>75%) is partitioned at finer scales (i.e., between samples within localities and within samples; Appendix S5) and rates of inbreeding are fairly high across its range ($F_{is} = 0.487$). Although *V. lantanoides* can reproduce through self-pollination, most fruit production occurs through outcrossing (Park et al., 2019), so the patterns observed here are most likely due to a combination of low gene flow and consanguineous mating within increasingly isolated populations. The results for *V. lantanoides* are likely to be representative of other ENA shrubs with similar distributions, and demonstrate the need for additional phylogeographic studies of such species to achieve a more complete understanding of the biogeographic history of ENA.

Survival where and how?

Our genetic and paleoclimatic analyses provide strong evidence for northward range expansion from a southern refugial area in *V. lantanoides* (Fig. 4B–D). However, models generated from different climatic reconstructions conflict somewhat with respect to the most likely locations of high-suitability climates that may have served as glacial refugia.

Both CCSM4 (Fig. 4B) and MPI (Fig. 4D) models identify a region of high climatic suitability in the Atlantic Coastal Plain along the coast of North Carolina. Paleoecological studies have reconstructed this region as having been inhabited by cold-temperate and boreal forest species (e.g., *Abies* and *Picea*) during the LGM (Jackson et al., 2000). A handful of phylogeographic studies have suggested that this area served as a glacial refugium for a number of temperate forest species (Morris et al., 2008; Barnard-Kubow et al., 2015; Nadeau et al., 2015; Bemmels and Dick, 2018). Considering the cold-temperate climatic niche of *V. lantanoides*, and the fact that it is found exclusively in spruce–fir outcrops at high elevations in the southern Appalachians, it is quite possible that *V. lantanoides* survived the LGM in this region. There are records of fossil *Viburnum* pollen in low frequency during the LGM in eastern North Carolina (Whitehead, 1981). However, it is not clear that these grains are from *V. lantanoides*, as opposed to other ENA *Viburnum* species.

By contrast, the MIROC model (Fig. 4C) identifies a large region of suitable climate in the Gulf Coastal Plain centered around the Lower Mississippi River Valley. A number of paleoecological studies have argued that this area served as a glacial refugium for temperate forest species, with cold-adapted species surviving inland, and less cold-tolerant species surviving along the Gulf Coast (Delcourt and Delcourt, 1984; Prentice et al., 1993; Jackson et al., 2000; Williams, 2003). Phylogeographic studies of several other plant groups have also suggested that deciduous and coniferous trees, understory shrubs, and herbs may have survived both in the interior and along the coast of the Gulf Coastal Plain (Griffin and Barrett, 2004b; Gonzales et al., 2008; Morris et al., 2010; Saeki et al., 2011; Peterson and Graves, 2015; Vargas-Rodriguez et al., 2015). *Viburnum* pollen grains dated to the LGM have been found in coastal North Carolina (Whitehead, 1981), throughout the Gulf Coastal Plain (Delcourt et al., 1980; Jackson and Givens, 1994), and as far north as the Upper Mississippi River Valley (Delcourt, 1979). Although these grains may have been produced by other ENA *Viburnum* species (Spriggs et al., unpublished data), it is certainly plausible that the interior of the Gulf Coastal Plain served as a refugium for *V. lantanoides*.

A third possibility is that *V. lantanoides* was more widespread during the LGM than suggested by these paleoclimate distribution models, possibly surviving in both refugial areas in the interior portions of the Atlantic and Gulf Coastal Plains, as well as in small pockets of suitable habitat scattered in between. This interpretation is consistent especially with our LGM species distribution models using the CCSM4 and MPI climate models (Fig. 4B, D). Paleoecological (Braun, 1955) and phylogeographic (Magni et al., 2005; McLachlan et al., 2005) studies have argued for the existence of glacial refugia in and around the southern Appalachians, where temperate forest species may have sought refuge in protected coves and ravines during glacial periods. Our inability to identify more suitable climates in this region is likely due to the fact that the spatial grain of our climatic data (2.5 arc-minutes, 5 km^2) is too coarse to capture microclimatic variation at fine spatial scales (Dobrowski, 2011; Gavin et al., 2014). However, given the presence of multiple, distinct genetic lineages in the southern Appalachians (Fig. 3), it is easy to imagine a scenario in which *V. lantanoides* was more widespread, with the modern southern Appalachian populations representing remnants of the broader LGM population.

One possibility that we can confidently rule out is that *V. lantanoides* survived the LGM in a cryptic refugium (a nunatak or driftless area) at higher latitudes (e.g., Godbout et al., 2010). It also seems likely that it did not survive along the southern coast, where many wide-ranging temperate forest species with more southerly modern range limits are thought to have survived the LGM (Jackson et al., 2000). Survival outside of traditionally circumscribed macrorefugia on the coastal plain may be common in other cold-adapted ENA shrub species. Paleoclimatic distribution models suggest that the eastern leatherwood, *Dirca paulustris*, survived in refugial areas distributed at midlatitudes (i.e., 35–40°N), spanning from the Ozark Plateau to coastal Virginia (Peterson and Graves, 2015). Furthermore, populations in the southern portion of its current range were found to be genetically depauperate when compared to populations at higher latitudes, providing additional evidence for glacial survival at midlatitudes (Peterson and Graves, 2015). Similarly, recent phylogeographic studies of ENA *Viburnum* species in the *Lentago* clade indicate that several of these species with more northern modern distributions (*V. lentago*, *V. prunifolium*, and *V. cassinoides*) are likely to have survived the LGM in midlatitudinal refugia (Spriggs et al., unpublished data). By contrast, *Lentago* clade species with more southern distributions in ENA (*V. nudum*, *V. nitidum*, *V. obovatum*, and *V. rufidulum*) are inferred to have occupied suitable habitats along the Gulf coastal plain and in Florida during the LGM (Spriggs et al., 2019). These findings illustrate the heterogeneity of responses to the LGM and provide new insights into plant movements in response to climate change in ENA.

Dispersal and diversification in ENA

By our current count, eastern North America is home to 17 species of *Viburnum* belonging to six widely separated clades: seven species of the *Lentago* clade; three of *Mollotinus*; three of *Dentata*; two of *Opulus*; one of *Lobata*; and one (*V. lantanoides*) of *Pseudotinus* (Clement and Donoghue, 2011; Spriggs et al., 2015; Eaton et al., 2017; M.J. Landis et al., unpublished data). In three of these lineages (*Lentago*, *Mollotinus*, and *Dentata*) speciation is inferred to have taken place within ENA. The four species of the other three clades (*V. opulus*, *V. edule*, *V. acerifolium*, and *V. lantanoides*) have broad northern geographic distributions but have not differentiated significantly across their ranges. As we have shown, *V. lantanoides* exhibits relatively little genetic diversity across its range. Likewise, it shows little consistent morphological differentiation in characters that distinguish species in the other ENA clades (i.e., leaf size and shape, leaf margins, and pubescence). And, whereas distinct differences in habitat and flowering time have evolved in other ENA *Viburnum* clades (e.g., within the *V. nudum* and *V. lentago* complexes; Spriggs et al., unpublished data), such differences are not evident in *V. lantanoides*, or in the other three widespread northern species. These observations beg the question: Why have some lineages diversified whereas others, including *V. lantanoides*, have not?

The answer may lie in the timing of dispersal into the New World, and a difference in the ancestral niches of the lineages that entered North America. Two major *Viburnum* lineages (*Lentago* and *Porphyrotinus*, the latter including both the *Mollotinus* and *Dentata* clades) appear to have entered North America much earlier than the others, most likely in the late Eocene or early Oligocene (Winkworth and Donoghue, 2005; Spriggs et al., 2015; M. J. Landis et al., unpublished data). Furthermore, these two major lineages are inferred to have occupied subtropical or warm-temperate forests at that time (Edwards et al., 2017; M. J. Landis et al., unpublished data). In these lineages, adaptations to colder climates are thought to have evolved later, especially as climates cooled in the Oligocene and Miocene. Ultimately, diversification in the *Lentago* and *Dentata* clades yielded separate species in the Deep South, across central North America, and in the Northeast.

In stark contrast to this pattern, the three other major lineages that entered North America (*V. edule* and *V. trilobum* in the *Opulus* clade, *V. acerifolium* in *Lobata*, and *V. lantanoides* in *Pseudotinus*) are inferred to have arrived more recently (from the mid-Miocene onward; Spriggs et al., 2015; M. J. Landis et al., unpublished data). Furthermore, the ancestors of these lineages were already adapted to cold temperate conditions when they arrived. *Viburnum lantanoides* clearly shows this second pattern. Its ancestors appear to have entered North America, probably through Beringia, during the middle or late Miocene (~12 mya; Spriggs et al., 2015; M. J. Landis et al., unpublished data) when colder, more seasonal climates were spreading around the Northern Hemisphere. The three close relatives of *V. lantanoides* in the *Pseudotinus* clade occupy cool-temperate forests in eastern Asia (*V. sympodiale* and *V. furcatum*) or more boreal forests in the Himalayas (*V. nervosum*), and therefore it is likely that the ancestral populations that gave rise to *V. lantanoides* were also well adapted to cold climates.

We hypothesize that as climates fluctuated from the mid-Miocene to the present, the lineages that entered North America earlier and in subtropical forests progressively diversified into colder climates. By contrast, the lineages that entered more recently, and which were already adapted to cold climates, did not diversify into warmer climates. This may fit a broader pattern in *Viburnum* in which there appear to have been multiple evolutionary shifts from warmer into colder climates, and few shifts in the other direction (Schmerler et al., 2012; Spriggs et al., 2015; Edwards et al., 2017). These observations highlight the value of placing phylogeographic analyses of particular species into a broader phylogenetic framework. The ways in which species have responded (or not) to climate changes, and their modern genetic and morphological diversity patterns, may, to a considerable extent, reflect their deeper evolutionary histories.

CONCLUSIONS

Our analyses paint the following picture for *V. lantanoides*. The ancestor of *V. lantanoides* had migrated into North America from Asia in the Miocene and was already adapted to cold climates. As glaciers receded following the LGM, *V. lantanoides* moved northward from a refugial area south of its current range. North of a boundary that roughly coincides with the southern edge of the last ice sheets, it appears that a single population of *V. lantanoides* rapidly expanded to the north and east. In the meantime, populations in the south became restricted to relict boreal forests at high elevations in the southern Appalachians. As a result of these processes, we now find significantly higher genetic diversity in the southern as compared to the northern part of its geographic range. We hypothesize that a similar pattern will be recovered in other patchily distributed species of shrubs and herbs in northeastern North America.

ACKNOWLEDGEMENTS

We thank the many state and provincial park and national forest employees who provided us with information on locating populations and granted us permission to collect. We also thank M. Sinnott-Armstrong, D. A. R. Eaton, R. I. Eytan, A. Ghezelayagh, D. J. MacGuigan, T. J. Near, N. P. Saarman, E. L. Spriggs, and two anonymous reviewers for their helpful comments on improving the analyses and/or manuscript. This work was funded in part by grants awarded to B.P. from the American Society of Plant Taxonomists and the Torrey Botanical Society.

AUTHOR CONTRIBUTIONS

B.P. and M.J.D. conceived the study and prepared the manuscript. B.P. conducted fieldwork and laboratory work and performed the analyses.

DATA ACCESSIBILITY

Demultiplexed sequence data are available on the NCBI Short Read Archive (https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP170982). Additional data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.1748225 (Park and Donoghue, 2018).

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

APPENDIX S1. Locality data for populations sampled in this study.

APPENDIX S2. Schematic depiction of the demographic models tested in this study.

APPENDIX S3. The approximate geographic range of *V. lantanoides*.

APPENDIX S4. Global population genetic parameter estimates for *V. lantanoides*.

APPENDIX S5. The distribution of genetic diversity at various levels of organization inferred through AMOVA.

LITERATURE CITED

Alberto, F. J., S. N. Aitken, R. Alía, S. C. González-Martínez, H. Hänninen, A. Kremer, F. Lefèvre, et al. 2013. Potential for evolutionary responses to climate change—evidence from tree populations. *Global Change Biology* 19: 1645–1661.

Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19: 1655–1664.

Bai, W. N., W. J. Liao, and D. Y. Zhang. 2010. Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. *New Phytologist* 188: 892–901.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One* 3: e3376.

Barbet-Massin, M., F. Jiguet, C. H. Albert, and W. Thuiller. 2012. Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution* 3: 327–338.

Barnard-Kubow, K. B., C. L. Debban, and L. F. Galloway. 2015. Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. *American Journal of Botany* 102: 1842–1853.

Bemmels, J. B., and C. W. Dick. 2018. Genomic evidence of a widespread southern distribution during the Last Glacial Maximum for two eastern North American hickory species. *Journal of Biogeography* 2018: 1739–1750.

Braun, E. L. 1955. The phytogeography of unglaciated eastern United States and its interpretation. *Botanical Review* 21: 297–375.

Broyles, S. 1998. Postglacial migration and the loss of allozyme variation in northern populations of *Asclepias exaltata* (Asclepiadaceae). *American Journal of Botany* 85: 1091.

Brunsfeld, S. J., J. Sullivan, D. E. Soltis, and P. S. Soltis. 2001. Comparative phylogeography of northwestern North America: A synthesis. *In* J. W. Silvertown [ed.], Integrating ecology and evolution in a spatial context, 319–339. Blackwell, Williston, Vermont, USA.

Burnham, K. P., and D. R. Anderson. 2007. Model selection and multimodel inference: A practical information-theoretic approach. Springer Science and Business Media, New York, New York, USA.

Caye, K., T. M. Deist, H. Martins, O. Michel, and O. François. 2016. TESS3: Fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources* 16: 540–548.

Chifman, J., and L. Kubatko. 2014. Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30: 3317–3324.

Clement, W. L., and M. J. Donoghue. 2011. Dissolution of *Viburnum* section *Megalotinus* (Adoxaceae) of Southeast Asia and its implications for morphological evolution and biogeography. *International Journal of Plant Sciences* 172: 559–573.

Coffman, A. J., P. H. Hsieh, S. Gravel, and R. N. Gutenkunst. 2016. Computationally efficient composite likelihood statistics for demographic inference. *Molecular Biology and Evolution* 33: 591–593.

Delcourt, H. R. 1979. Late Quaternary vegetation history of the eastern Highland Rim and adjacent Cumberland Plateau of Tennessee. *Ecological Monographs* 49: 255–280.

Delcourt, H. R., and P. A. Delcourt. 1984. Ice age haven for hardwoods. *Natural History* 93: 22–28.

Delcourt, P. A., H. R. Delcourt, R. C. Brister, and L. E. Lackey. 1980. Quaternary vegetation history of the Mississippi Embayment. *Quaternary Research* 13: 111–132.

Dobrowski, S. Z. 2011. A climatic basis for microrefugia: The influence of terrain on climate. *Global Change Biology* 17: 1022–1035.

Dray, S., and A. B. Dufour. 2007. The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software* 22.

Eaton, D. A. R. 2014. PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30: 1844–1849.

Eaton, D. A. R., E. L. Spriggs, B. Park, and M. J. Donoghue. 2017. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology* 66: 399–412.

Edwards, E. J., D. S. Chatelet, B. C. Chen, J. Y. Ong, S. Tagane, H. Kanemistu, K. Tagawa, et al. 2017. Convergence, consilience, and the evolution of temperate deciduous forests. *The American Naturalist* 190: S87–S104.

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

Frichot, E., F. Mathieu, T. Trouillon, G. Bouchard, and O. François. 2014. Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196: 973–983.

Gavin, D. G., M. C. Fitzpatrick, P. F. Gugger, K. D. Heath, F. Rodríguez-Sánchez, S. Z. Dobrowski, A. Hampe, et al. 2014. Climate refugia: Joint inference from fossil records, species distribution models and phylogeography. *New Phytologist* 204: 37–54.

Godbout, J., J. Beaulieu, and J. Bousquet. 2010. Phylogeographic structure of jack pine (*Pinus banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern North America. *American Journal of Botany* 97: 1903–1912.

Gonzales, E., J. L. Hamrick, and S. M. Chang. 2008. Identification of glacial refugia in south-eastern North America by phylogeographical analyses of a forest understorey plant, *Trillium cuneatum*. *Journal of Biogeography* 35: 844–852.

Gosselin, T. 2017. radiator: RADseq data exploration, manipulation and visualization using R. https://github.com/thierrygosselin/radiator.

Gosselin, T., E. C. Anderson, and A. L. Ferchaud. 2016. Assigner: Population assignment analysis of genomic dataset (e.g. GBS/RADseq) using R. https://github.com/thierrygosselin/assigner.

Goudet, J. 2005. hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Resources* 5: 184–186.

Gould, W. P. 1966. The ecology of *Viburnum alnifolium* Marsh. Ph.D. dissertation, State University College of Forestry at Syracuse University.

Griffin, S. R., and S. C. H. Barrett. 2004a. Genetic variation in *Trillium erectum* (Melanthiaceae), a widespread forest herb in eastern North America. *Canadian Journal of Botany* 82: 316–321.

Griffin, S. R., and S. C. H. Barrett. 2004b. Post-glacial history of *Trillium grandiflorum* (Melanthiaceae) in eastern North America: Inferences from phylogeography. *American Journal of Botany* 91: 465–473.

Gutenkunst, R. N., R. D. Hernandez, S. H. Williamson, and C. D. Bustamante. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics* 5: e1000695.

Healey, A., A. Furtado, T. Cooper, and R. J. Henry. 2014. Protocol: A simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. *Plant Methods* 10: 21.

Hewitt, G. M. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68: 87–112.

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

Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Proceedings of the Royal Society B* 359: 183–195.

Hijmans, R. J. 2016. Geosphere: Spherical trigonometry. R package version 1.5-5.

Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith. 2017. dismo: Species distribution modeling. R Package version 1.1–4.

Hu, L. J., K. Uchiyama, H. L. Shen, Y. Saito, Y. Tsuda, and Y. Ide. 2008. Nuclear DNA microsatellites reveal genetic variation but a lack of phylogeographical structure in an endangered species, *Fraxinus mandshurica*, across north-east China. *Annals of Botany* 102: 195–205.

Huang, H., and L. Knowles. 2014. Unforeseen consequences of excluding missing data from next-generation sequences: Simulation study of RAD sequences. *Systematic Biology* 65: 357–365.

Ibrahim, K. M., R. A. Nichols, and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* 77: 282–291.

Ingvarsson, P. K. 2008. Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula*. *Genetics* 180: 329–340.

Jackson, S. T., and C. R. Givens. 1994. Late Wisconsinan vegetation and environment of the Tunica Hills Region, Louisiana/Mississippi. *Quaternary Research* 41: 316–325.

Jackson, S. T., R. S. Webb, K. H. Anderson, J. T. Overpeck, T. Webb, J. W. Williams, and B. C. S. Hansen. 2000. Vegetation and environment in eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews* 19: 489–508.

Jaramillo-Correa, J. P., J. Beaulieu, D. P. Khasa, and J. Bousquet. 2009. Inferring the past from the present phylogeographic structure of North American forest trees: Seeing the forest for the genes. *Canadian Journal of Forest Research* 39: 286–307.

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

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics* 11: 94.

Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2: e281.

Krutovsky, K. V., J. Burczyk, I. Chybicki, R. Finkeldey, T. Pyhäjärvi, and J. J. Robledo-Arnuncio. 2012. Gene flow, spatial structure, local adaptation, and assisted migration in trees. *In* R. J. Schnell and P. M. Priyadarshan [eds.], Genomics of tree crops, 71–116. Springer, Dordrecht, The Netherlands.

Lewis, P. O., and D. J. Crawford. 1995. Pleistocene refugium endemics exhibit greater allozymic diversity than widespread congeners in the genus *Polygonella* (Polygonaceae). *American Journal of Botany* 82: 141.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079.

Lumibao, C. Y., S. M. Hoban, and J. McLachlan. 2017. Ice ages leave genetic diversity 'hotspots' in Europe but not in Eastern North America. *Ecology Letters* 20: 1459–1468.

Magni, C. R., A. Ducousso, H. Caron, R. J. Petit, and A. Kremer. 2005. Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. *Molecular Ecology* 14: 513–524.

McLachlan, J. S., J. S. Clark, and P. S. Manos. 2005. Molecular indicators of tree migration capacity under rapid climate change. *Ecology* 86: 2088–2098.

Morris, A. B., S. M. Ickert-Bond, D. B. Brunson, D. E. Soltis, and P. S. Soltis. 2008. Phylogeographical structure and temporal complexity in American sweetgum (*Liquidambar styraciflua*; Altingiaceae). *Molecular Ecology* 17: 3889–3900.

Morris, A. B., C. H. Graham, D. E. Soltis, and P. S. Soltis. 2010. Reassessment of phylogeographical structure in an eastern North American tree using Monmonier's algorithm and ecological niche modelling. *Journal of Biogeography* 37: 1657–1667.

Nadeau, S., J. Godbout, M. Lamothe, M. C. Gros-Louis, N. Isabel, and K. Ritland. 2015. Contrasting patterns of genetic diversity across the ranges of *Pinus monticola* and *P. strobus*: A comparison between eastern and western North American postglacial colonization histories. *American Journal of Botany* 102: 1342–1355.

Park, B., and M. J. Donoghue. 2018. Data from: Phylogeography of a widespread eastern North American shrub, *Viburnum lantanoides*. Dryad Digital Repository. https://doi.org/10.5061/dryad.1748225.

Park, B., M. Sinnott-Armstrong, C. Schlutius, J. P. Zuluaga, E. L. Spriggs, R. G. Simpson, E. Benavides, et al. 2019. Sterile marginal flowers increase visitation and fruit set in the hobblebush (*Viburnum lantanoides*, Adoxaceae) at multiple spatial scales. *Annals of Botany* 123: 381–390.

Peirson, J. A., C. W. Dick, and A. A. Reznicek. 2013. Phylogeography and polyploid evolution of North American goldenrods (*Solidago* subsect. *Humiles*, Asteraceae). *Journal of Biogeography* 40: 1887–1898.

Peterson, B. J., and W. R. Graves. 2015. Chloroplast phylogeography of *Dirca palustris* L. indicates populations near the glacial boundary at the Last Glacial Maximum in eastern North America. *Journal of Biogeography* 43: 314–327.

Petit, R. J. 2003. Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* 300: 1563–1565.

Phillips, S. J., M. Dudík, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. *In* Twenty-First International Conference on Machine Learning—ICML '04.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259.

Prentice, I. C., P. J. Bartlein, and T. Webb. 1993. Vegetation and climate change in eastern North America eince the Last Glacial Maximum. *Ecology* 74: 998.

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

Qiu, Y. X., C. X. Fu, and H. P. Comes. 2011. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution* 59: 225–244.

Rubin, B. E. R., R. H. Ree, and C. S. Moreau. 2012. Inferring phylogenies from RAD sequence data. *PLoS One* 7: e33394.

Saeki, I., C. W. Dick, B. V. Barnes, and N. Murakami. 2011. Comparative phylogeography of red maple (*Acer rubrum* L.) and silver maple (*Acer saccharinum* L.): Impacts of habitat specialization, hybridization and glacial history. *Journal of Biogeography* 38: 992–1005.

Sage, R. D., and J. O. Wolff. 1986. Pleistocene glaciations, fluctuating ranges, and low genetic variability in a large mammal (*Ovis dalli*). *Evolution* 40: 1092–1095.

Savolainen, O., T. Pyhäjärvi, and T. Knürr. 2007. Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics* 38: 595–619.

Schmerler, S. B., W. L. Clement, J. M. Beaulieu, D. S. Chatelet, L. Sack, M. J. Donoghue, and E. J. Edwards. 2012. Evolution of leaf form correlates with tropical–temperate transitions in *Viburnum* (Adoxaceae). *Proceedings of the Royal Society B: Biological Sciences* 279: 3905–3913.

Schwaegerle, K. E., and B. A. Schaal. 1979. Genetic variability and founder effect in the pitcher plant *Sarracenia purpurea* L. *Evolution* 33: 1210–1218.

Shafer, A. B. A., C. I. Cullingham, S. D. Côté, and D. W. Coltman. 2010. Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* 19: 4589–4621.

Spriggs, E. L., W. L. Clement, P. W. Sweeney, S. Madriñán, E. J. Edwards, and M. J. Donoghue. 2015. Temperate radiations and dying embers of a tropical past: The diversification of *Viburnum*. *New Phytologist* 207: 340–354.

Spriggs, E. L., D. A. R. Eaton, P. W. Sweeney, C. Schlutius, E. J. Edwards, and M. J. Donoghue. 2019. Restriction-Site-Associated DNA sequencing reveals a cryptic *Viburnum* species on the North American Coastal Plain. *Systematic Biology* 68: 187–203.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.

Suarez-Gonzalez, A., J. T. Sutton, A. J. Trant, E. Zamlynny, and S. V. Good. 2015. Rethinking refugia: Tree topology, divergence dates, and demographic history trace the distribution of the endangered Plymouth gentian (*Sabatia kennedyana*) from the Pleistocene glaciation to present day. *American Journal of Botany* 102: 609–620.

Thomson, A. M., C. W. Dick, and S. Dayanandan. 2014. A similar phylogeographical structure among sympatric North American birches (*Betula*) is better explained by introgression than by shared biogeographical history. *Journal of Biogeography* 42: 339–350.

Tian, B., R. Liu, L. Wang, Q. Qiu, K. Chen, and J. Liu. 2009. Phylogeographic analyses suggest that a deciduous species (*Ostryopsis davidiana* Decne., Betulaceae) survived in northern China during the Last Glacial Maximum. *Journal of Biogeography* 36: 2148–2155.

Varela, S., M. S. Lima-Ribeiro, and L. C. Terribile. 2015. A short guide to the climatic variables of the Last Glacial Maximum for biogeographers. *PLoS One* 10: e0129037.

Vargas-Rodriguez, Y. L., W. J. Platt, L. E. Urbatsch, and D. W. Foltz. 2015. Large scale patterns of genetic variation and differentiation in sugar maple from tropical Central America to temperate North America. *BMC Evolutionary Biology* 15: 257.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.

Whitehead, D. R. 1981. Late-Pleistocene vegetational changes in northeastern North Carolina. *Ecological Monographs* 51: 451–471.

Williams, J. W. 2003. Variations in tree cover in North America since the last glacial maximum. *Global and Planetary Change* 35: 1–23.

Williams, J. W., B. N. Shuman, T. Webb, P. J. Bartlein, and P. L. Leduc. 2004. Late-Quaternary vegetation dynamics in North America: Scaling from taxa to biomes. *Ecological Monographs* 74: 309–334.

Winkworth, R. C., and M. J. Donoghue. 2005. *Viburnum* phylogeny based on combined molecular data: Implications for taxonomy and biogeography. *American Journal of Botany* 92: 653–666.

Wright, S. 1943. Isolation by distance. *Genetics* 28: 114–138.

Zeng, Y. F., W. T. Wang, W. J. Liao, H. F. Wang, and D. Y. Zhang. 2015. Multiple glacial refugia for cool-temperate deciduous trees in northern East Asia: The Mongolian oak as a case study. *Molecular Ecology* 24: 5676–5691.

Zinck, J. W. R., and O. P. Rajora. 2016. Post-glacial phylogeography and evolution of a wide-ranging highly-exploited keystone forest tree, eastern white pine (*Pinus strobus*) in North America: Single refugium, multiple routes. *BMC Evolutionary Biology* 16: 56.