Biogeography and ecological niche evolution in Diapensiaceae inferred from phylogenetic analysis

Michelle L. Gaynor1,2*, Chao-Nan Fu3, Lian-Ming Gao3, Li-Min Lu4, Douglas E. Soltis1,2 and Pamela S. Soltis1

1 Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA
2 Department of Biology, University of Florida, Gainesville, FL 32611, USA
3 CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China
4 State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
*Author for correspondence. E-mail: michellegaynor@ufl.edu
Received 24 November 2019; Accepted 29 May 2020; Article first published online 5 June 2020

Abstract Diapensiaceae (Ericales) are a small family of about 15 species. Within this clade, two species are broadly distributed throughout the Northern Hemisphere, whereas the remaining species have a disjunct distribution between eastern North America and eastern Asia. To address patterns and processes of diversification in Diapensiaceae, we conducted biogeographic analyses and inferred shifts in the ecological niche across the phylogeny of the clade. Although Diapensiaceae have been the focus of multiple phylogenetic and biogeographic studies, previous studies have been taxonomically limited. This study has greatly improved the phylogenetic underpinning for Diapensiaceae with the most inclusive taxonomic sampling thus far, employing both nuclear and plastid gene sequence data for at least one sample per species in the family. Our estimates indicate that genera of Diapensiaceae variously diverged in the Eocene, Oligocene, and early to mid-Miocene. The biogeographic analysis suggests that the probable ancestor of the Diapensiaceae crown clade originated in the Nearctic, with vicariance events contributing to the current distribution of the disjunct taxa. Ecological niche, when considered in a phylogenetic context, was observed to be clustered on the basis of biogeographic realm. In general, a greater ecological overlap was found at younger nodes and a greater niche divergence was found among distantly related species. Diversification in Diapensiaceae appears to have been shaped by both large-scale biogeographic factors, such as vicariance, and divergence in an ecological niche among closely related species. Key words: Angiosperm-353 probes, comparative phylogenetics, disjunction, ecological niche divergence, target enrichment.

1 Introduction

Floristic disjunctions across the Northern Hemisphere have long captivated the attention of evolutionary botanists (Gray, 1859; Li, 1952; Graham, 1966; Boufford & Spongberg, 1983; Xiang & Soltis, 2001; Donoghue & Smith, 2004; Manos & Meireles, 2015). Many plant groups exhibit the well-known disjunction between eastern North America (ENA) and eastern Asia (EA; reviewed in Hong, 1993; Xiang et al., 1998; Qian & Ricklefs, 1999; Wen, 1999; Donoghue & Smith, 2004), and migration pathways, timing, and the biogeographic region of origin have been the focus of many studies of these disjunct lineages (reviewed by Wen et al., 2006 & Graham, 2008). However, the underlying processes that shaped this disjunction are unclear and may vary among clades, as suggested by incongruent age estimates for the disjunction among taxa (e.g., Xiang et al., 1998; Manos & Meireles, 2015).

The ENA–EA disjunction has been related to geological and climatic changes during the Cenozoic era, which shaped the distribution of the Boreal tropical flora in the Northern Hemisphere (Wolfe, 1975; Tiffney, 1985a; Wen, 1999; Manos & Meireles, 2015). In addition to vicariance and biotic interactions as drivers of this disjunction, the spread of taxa in the Northern Hemisphere is believed to have been enabled by migration or long-distance dispersal via land bridges, specifically the North Atlantic Land Bridge and the Bering Land Bridge (Tiffney, 1985a, 1985b; Sanmartín et al., 2001; Tiffney & Manchester, 2001; Donoghue & Smith, 2004; Wen et al., 2010). Thus, many processes, from large-scale vicariance to shifts in an ecological niche, likely contributed to patterns of diversification in clades exhibiting this disjunction. However, analyses of additional clades are needed to identify the extent to which alternative processes may have contributed to current distributional patterns.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
The identification of the historical processes that shaped the ENA–EA disjunction in any clade requires a robust, dated phylogenetic framework, biogeographic distributions, and knowledge (or inference) of the ecological niches of the species. Insight into species interactions can be assessed by understanding both the realized and fundamental niche in a phylogenetic context. A species’ realized niche encompasses abiotic conditions that a species can occupy with the presence of biotic interactions, whereas a species’ fundamental niche comprises the abiotic conditions a species could potentially occupy in the absence of biotic interactions (Peterson et al., 2011). Ecological niche can be examined among clades to identify phylogenetic niche conservation or phylogenetic niche divergence, respectively the retention or differentiation of an ancestral niche (Harvey & Pagel, 1991; Wiens & Graham, 2005; Pyron et al., 2015). In addition, modes of speciation can be inferred on the basis of niche space considered in a phylogenetic context; allopatric speciation is often related to phylogenetic niche conservation, whereas parapatric or sympatric speciation may result from phylogenetic niche divergence (Fitzpatrick & Turelli, 2006; Cardillo & Warren, 2016).

Diapensiaceae are a small family in Ericales (APG IV et al., 2016), comprising six genera and, until recently, 15 species that are all distributed in the Northern Hemisphere (Fig. S1). Diapensiaceae belong to a small clade of families, along with Styraceae and Symphoricarpos, referred to as the “styracoids” (Schönberger et al., 2005). In Diapensiaceae, Galax Sims and Berneuxia Decne. are both monotypic, and Pyxidanthera Michx., Shortia Torrey & A. Gray, Diapensia L., and Schizocodon Siebold and Zucc. include, respectively, two species, four species, five species, and two species. Two additional species of Shortia have been recently described (see Section 2).

The majority of the species in Diapensiaceae are distributed in EA and ENA, and their restriction to these centers of diversity may reflect relictual ranges of a once more widespread distribution in the Northern Hemisphere (Scott, 2004). Galax urceolata (Poir.) Brummitt, Pyxidanthera barbula Michx., P. brevifolia Wells, and Shortia galacifolia Torr. & A. Gray are all endemic to ENA. Shortia rotundifolia (Maxim.) Makino occurs in Taiwan, and S. sinensis Hemsley., Berneuxia thibetica Decaisne, Diapensia purpurea Diels, D. wardii W. E. Evans, and D. himalaica J. D. Hooker & Thomon all have distributions centered in southwestern China (some ranging from Tibet into India, Nepal, and Vietnam). Both species of Schizocodon, Sc. soldanelloides Siebold & Zucc. and Sc. ilicifolius Maxim., as well as Shortia uniflora Maxim., are endemic to the Japanese archipelago. Although most species in Diapensiaceae have distributions in either EA or ENA, two species have broader distributions in the Northern Hemisphere. Diapensia obovata (F. Schmidt) Nakai has an amph-Beringian distribution, and D. laponica L. has an amph-Atlantic distribution (Hultén, 1958; Scott & Day, 1983; Malychev, 1997; Elven et al., 2011; Hou et al., 2016), both have been referred to as “mainly arctic” (Hultén & Fries, 1986). The biogeographic and ecological factors that shaped the current distribution of Diapensiaceae across EA and ENA, as well as the timing of divergence among these taxa, are not known.

To identify the factors shaping the distribution of Diapensiaceae, including biogeographic and ecological evolution, the current phylogenetic underpinning needs improvement. Previous phylogenetic studies of Diapensiaceae have been taxonomically limited and have found conflicting relationships among genera. The most inclusive phylogenetic analysis to date (Rönblom & Anderberg, 2002) was based on morphology and four genes (one nuclear, two plastid, and one mitochondrial), but it included molecular data for only eight species: B. thibetica, D. laponica, G. urceolata, P. barbula, S. rotundifolia (synonym S. exappendiculata), S. galacifolia, S. uniflora, and Sc. soldanelloides (as Shortia soldanelloides (Sieb. & Zucc.) Makino). The lack of complete taxonomic sampling may confound the inferred phylogenetic relationships. In previous studies, the North American Galax was sister to all other genera (Rönblom & Anderberg, 2002; Higashi et al., 2015), whereas Pyxidanthera, also from North America, was the subsequent sister to the remaining genera (Rönblom & Anderberg, 2002). Notably, the circumscription and relationships of Shortia differed in previous studies. Rönblom & Anderberg (2002) found that Schizocodon soldanelloides (recognized as Shortia soldanelloides) formed a clade with species of Shortia, but relationships between Shortia and the remaining genera were unresolved. In contrast, Higashi et al. (2015), using multiple nuclear markers and a subset of taxa, found Sc. soldanelloides and Sc. ilicifolius to be sisters and most closely related to D. laponica, rather than the sister to Shortia. On the basis of these two studies, other than the placements of Galax and Pyxidanthera, the relationships among genera of Diapensiaceae remain uncertain.

Within genera with two or more species, a clear understanding of relationships is similarly lacking. On the basis of morphological and genetic data, there is a taxonomic uncertainty regarding the distinction between P. barbula and P. brevifolia (Wall et al., 2010). Previous studies also suggest, based on AFLPs, that multiple introgression events occurred between Sc. soldanelloides and Sc. ilicifolius due to range shifts during the Pleistocene (Higashi et al., 2013). Within Diapensia, the two species found in the Himalayan–Hengduan Mountains, D. purpurea and D. himalaica, were identified as sisters (Hou et al., 2015), and the broadly distributed D. obovata and D. laponica were also sisters to each other and together formed a sister group with D. purpurea and D. himalaica (Hou et al., 2015). However, no studies to date have included D. wardii. Shortia rotundifolia and S. uniflora were identified as most closely related to each other and sister to S. galacifolia; together these species are believed to be sister to S. sinensis, but more work is needed (Higashi et al., 2015).

For species of Diapensiaceae, dispersal is thought to be limited, as seed germination occurs while the seeds are retained in the capsule, and there is no obvious mechanism of dispersal (Ross, 1936; Primack & Wyatt, 1975; Scott, 2004). However, seed dispersal has rarely been studied in the family. Ants have been observed transporting seeds in Pyxidanthera (Wall et al., 2010); the dispersal distance is unknown and expected to be limited. Despite barriers to seed dispersal, capsules of Diapensia laponica have been observed to be dispersed by wind over snow and ice, which may have allowed range expansion (Day & Scott, 1981). Species examined within the family also have low pollen availability and deposition, which may limit gene flow (Elberling, 2001; Barringer & Galloway, 2017).

Limited dispersal abilities and the current distribution of some congeneric species have spurred interest in the biogeographic history of Diapensiaceae. In addition, due to their intercontinental disjunction, Diapensiaceae have been considered as an old lineage (Baldwin, 1939). Hou et al. (2016),
based on a study of Diapensia, estimated that the family originated approximately 65 Mya with a Northern Hemisphere origin, specifically in the Himalayan–Hengduan Mountains, Japan, or the Arctic. Rose et al. (2018) found that the most probable origin of crown Diapensiaceae was in the Nearctic, with a slightly younger estimated age of 59.9 Mya. Shortia galacifolia, found in North America, was inferred to be the result of a vicariance event from a Nearctic or Palearctic ancestor, dated at about 12 Mya (Rose et al., 2018). Higashi et al. (2015) similarly inferred that S. galacifolia diverged following a vicariance event, most likely from the ancestor of S. uniflora and S. rotundifolia, whereas the ancestor of Schizocodon, S. uniflora, and S. rotundifolia was inferred to have originated in the Japanese archipelago. The limited dispersal abilities of some species suggest that long-distance dispersal is unlikely.

We aimed to improve the phylogenetic underpinning for Diapensiaceae by including nearly all currently recognized species and sampling numerous plastid and nuclear loci. On the basis of this inclusive phylogeny, we infer the biogeographic history of the family as well as the timing of major divergences and disjunctions. We also examine patterns of ecological niche evolution in Diapensiaceae. Specifically, we infer the mode of speciation among close relatives and determine the roles of niche conservatism and niche divergence. Finally, we consider the contributions of biogeography and ecological niche divergence in shaping the current distributions of species of Diapensiaceae.

2 Material and Methods

2.1 Samples and sequencing for phylogenetic analysis

In all, 26 samples from specimens of 14 of the 15 species of Diapensiaceae (excluding Diapensia wardii) recognized before 2017 were obtained from the following herbaria: CAS, CM, FLAS, IDS, MO, and NCU (Table S1; Fig. S2). We did not include two species, Shortia rotata Gaddy & Nuraliev (Gaddy & Nuraliev, 2017) and Shortia brevistyla (P. A. Davies) Gaddy (Gaddy et al., 2019), described within the last 3 years due to the lack of specimens, DNA samples, and occurrence records. Additionally, subspecies were not specified due to the inability to discriminate occurrence records among subspecies. Approximately 1 cm² of leaf tissue or less was sampled from one to three specimens per species. Genomic DNA was extracted using a modified CTAB protocol (Doyle & Doyle, 1987). To remove excess polysaccharides, a HEPES buffer wash (Setoguchi & Ohba, 1995; Shepherd & McLay, 2011) was used for samples of Pyxidanthera, Schizocodon, and Shortia. All DNA extractions were run on a 1.2% agarose gel to visually assess DNA quality. Samples were also quantified using a Qubit® 3.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Double-barcoded Illumina libraries were constructed, and hybridization to the Angiosperms-353 probe set (Johnson et al., 2018; available from Arbor BioScience, Ann Arbor, MI, USA) and sequencing (150-bp paired-end reads) on an Illumina HiSeq instrument were conducted by Rapid Genomics (Gainesville, FL, USA). Off-target plastid sequences were also obtained. For four species (D. obovata, D. purpurea, S. uniflora, and S. rotundifolia), where only a single specimen was available per species, we obtained duplicate DNA extractions and sequences to ensure inclusion in the phylogeny. Sequences are available in the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) (SRR11282173–SRR11282192). We obtained sequences for the outgroup, Cyrilla racemiflora (Cyrillaceae) (Accession #: SRX4321588), from NCBI SRA using the SRA toolkit.

Assembled and raw plastid sequences for Berneuxia thibetica, Shortia sinensis, Diapensia hidalguensis, D. purpurea, and D. wardii were provided by Gao LM & Fu CN (unpublished data). Libraries for genome skimming were prepared by BGI (Shenzhen, China) following the manufacturer’s protocol. Shortia sinensis was sequenced using an Illumina HiSeq 2500, and all other plastid sequences were obtained using Illumina HiSeq X-ten instrument.

2.2 Assembly of gene sequences

Before processing reads, if an individual was sequenced more than once, we combined the sequencing reads from the same individual. Sequencing reads were trimmed and cleaned using Trimomatic (Bolger et al., 2014) through the SECAPR bioinformatic pipeline (Andermann et al., 2018) with a simple clip threshold of 5 bp, a palindromic clip threshold of 20 bp, seed mismatch threshold of 5 bp, and a head crop of 10 bp. HybPiper (Johnson et al., 2016) was used to recover target sequences from the Angiosperms-353 probe set (Johnson et al., 2018) and a plastid gene set extracted from Alniphyllum eberhardtii (Styracaceae; KX765434); there is no fully assembled plastid genome reported for any species in Diapensiaceae. Reads were mapped to de-gapped medoid sequences using BWA (Li & Durbin, 2009), and each gene was assembled de novo in SPAdes (Bankevich et al., 2012). Assemblies were then assessed using get_seq_lengths.py and hybpiper_stats.py. Gene sequences were then retrieved using retrieve_sequences.py.

2.3 Alignment and phylogenetic analysis

Before alignment, sequences were checked to make sure they were in the proper direction using Adjust_Direction.py from the python toolkit SuperCRUNCH v1.1 (Portik & Wiens, 2019). Mafft v7.607 (Katoh et al., 2002) was used to conduct a Smith–Waterman local alignment with Gotoch affine gap cost (–maxiterate 1000). Alignments were then refined with trimal v1.2 (Capella-Gutierrez et al., 2009) with a gap threshold of 0.5. After alignment, we assessed alignment statistics with seqkit (Shen et al., 2016) to limit the amount of missing data while still retaining genes shared among distantly related taxa, we manually filtered genes to only include genes found in 50% of individuals, similar to methods applied in Couvreur et al. (2019).

The retained gene alignments were concatenated with catbastazphyml.pl. Partitioning schemes were identified using PartitionFinder v2.1.1 (Lanfear et al., 2012), a greedy algorithm was used, and models were selected on the basis of the Bayesian information criterion score. To construct the maximum likelihood phylogeny, we used RAxML v8.0.0 (Stamatakis, 2006) under a single model, GTR+GAMMA, and 1000 bootstrap replicates to evaluate clade support for data sets including all genes, only nuclear genes, and only plastid genes. A low bootstrap support was found for conflicting topologies in phylogenetic inferences based on nuclear and plastid only data sets. To accommodate a possible variation among individual nuclear gene trees and to investigate
whether gene trees differ in topology from the species tree, we generated unrooted gene trees with RAxML, where each gene tree was constructed under a GTRGAMMA model with 100 bootstrap replicates (Soltis, 1998). We used newickutils (Junier & Zdobnov, 2010) to collapse nodes that had less than 50% bootstrap support (BS). Next, for the nuclear data set, we inferred the coalescent tree using ASTRAL-III (Zhang et al., 2018) with default settings. A quartet support, or the percentage of quartets in gene trees that agree with a branch, was generated (t 1). To assess if gene tree topologies were unique, conflicting, or concordant, we implemented phyparts (Smith et al., 2015) and visualized this output using phypartspiecharts.py (https://github.com/mossmatters/phyloscripts/tree/master/phypartspiecharts). To implement phyparts, all input trees for the ASTRAL analysis must be rooted; therefore, gene trees that did not contain the outgroup—because the gene was not captured for the outgroup—were not included. This analysis was, therefore, conducted with 31 nuclear gene trees, and these gene trees were rooted with the outgroup Cyrilla racemiflora using Phyx (Brown et al., 2017).

Due to the low number of genes included in the multispecies coalescence analyses, the resulting phylogeny may not represent the best tree across the entire nuclear genome for these lineages. Plastid genes are generally inherited uniparentally (but see Corriveau & Coleman, 1988; Hansen et al., 2007; Weihe et al., 2009); therefore, they violate the assumptions of multispecies coalescence methods. Hence, they were not included in this analysis. Some recent studies have proposed that plastid genes may behave in a similar manner as nuclear regions, suggesting these coalescence methods may be justified (Hansen et al., 2007; Walker et al., 2019; Stull et al., 2020). However, the use of this method on plastid genes is still controversial due to an incomplete understanding of plastid gene linkage (reviewed in Gonçalves et al., 2019). Therefore, we did not conduct coalescence analyses on our plastid gene set.

2.4 Divergence time estimation

We used a reduced sample set, with only one individual per species, to construct a maximum likelihood phylogeny from the nuclear and plastid concatenated matrix for dating inference. We chose to combine nuclear and plastid genes to enable the inclusion of species for which only plastid sequences are available. Furthermore, the conflict found between the separate nuclear and plastid data sets was minimal and represented soft incongruence (Seelanan et al., 1997; Johnson 2018). We estimated divergence times using a penalized likelihood approach (Sanderson, 2002) implemented with TreePL (Smith & O’Meara, 2012). We set the maximum age and minimum age of the root as 110.1 Mya and 101.6 Mya, respectively. We also constrained the Diapensiaceae crown node with maximum and minimum ages of 65 Mya and 59.9 Mya, respectively. The age chosen here to constrain the crown node is corroborated by the age of the fossil Actinocalyx bohrii Friis (dated at 85.8–70.6 Mya), which has been assigned to the stem of Diapensiaceae and Styracaceae in previous analyses (Friis, 1985; Rose et al., 2018). Smoothing parameters were determined through cross-validation. The resulting 1000 bootstrap trees were summarized with TreeAnnotator v.2.6.0 (Bouckaert et al., 2019) to obtain the maximum clade credibility chronogram with the corresponding node heights and confidence intervals.

2.5 Species occurrences

Species occurrences were collected to define the biogeographic regions and conduct ecological niche analysis, as described below. Using the R package spocc (Chamberlain et al., 2018), occurrence records were obtained from iDigBio (https://www.idigbio.org), GBIF (https://www.gbif.org/), and BISON (https://bison.usgs.gov) for all species in Diapensiaceae. Additional records were obtained from the Chinese Virtual Herbarium (http://www.cvh.ac.cn/). Due to the tendency of citizen scientists to identify Diapensia lapponica incorrectly (MacKenzie et al., 2017), we excluded occurrence records from GBIF for this species and retained specimen records from iDigBio. Then, all records were manually inspected and visualized to identify outliers and unlikely occurrences, as well as taxonomic synonyms. Additionally, occurrence records were removed that corresponded to an institution (including herbaria, botanical gardens, etc.), based on coordinates provided through CoordinateCleaner (Zika et al., 2019). We rounded latitude and longitude to the nearest hundredth, and then duplicate coordinates were removed. Improbable records were also removed; for example, we removed points that occurred on a different continent than the known range of a species. The final data set included 2552 occurrence points (Table S2). Despite our attempt to obtain additional occurrence records for D. wardii, only three records were obtained due to its very limited distribution; therefore, we were unable to construct an ecological niche model for this species. Both raw and cleaned occurrence records are publicly available (https://doi.org/10.5061/dryad.rn8pk0p6h).

2.6 Biogeographical analyses

A biogeographical reconstruction of Diapensiaceae was conducted using BioGeoBears (Matzke, 2013). Areas were defined on the basis of eight biogeographic realms as proposed by Olson et al. (2001), available from the World Wildlife Fund website (http://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world). Following Rose et al. (2018), we condensed these eight realms into six areas, combining Australasia and Oceania and excluding Antarctica: (A) Afrotop, (B) Australasia + Oceania, (C) Indo-Malaysia, (D) Nearctic, (E) Neotropics, and...
2.7 Niche evolution

We downloaded 19 bioclimatic layers from WorldClim at 2.5 arc minute resolution (Hijmans et al., 2005) for the current period. A convex hull was constructed around all occurrence points with a buffer of 0.2 degrees around each occurrence point. We then cropped the defined extent using the R package raster v.2.9.5 (Hijmans et al., 2013). Highly correlated layers were identified using a Pairwise Pearson correlation among the 19 cropped layers. When a pair had a Pearson correlation >0.75, one of the two variables was removed; the remaining layers were used in all subsequent analyses. The remaining seven bioclimatic layers were annual mean temperature (bio1), mean diurnal range (bio2), temperature annual range (bio7), mean temperature of wettest quarter (bio8), annual precipitation (bio12), precipitation seasonality (bio15), and precipitation of coldest quarter (bio19). These variables were then cropped for each species using a convex hull constructed on each species’ occurrence points and a buffer of 0.2 degrees, using the methods described above.

First, we investigated the realized niche of all species by point sampling the current bioclimatic niche at each occurrence record. We then conducted a principal component analysis (PCA) in R version 3.2.1 (R Development Core Team, 2015). For each climatic variable, we analyzed niches among all included species using an analysis of variance (ANOVA), followed by post-hoc tests using the Tukey’s honest significant differences (TukeyHSD) method in R.

MaxEnt v.3.3.3k (Phillips et al., 2004, 2006) was used to generate an ecological niche model, or species distribution model, for each species on the basis of the seven uncorrelated bioclimatic variables and a maximum of 5000 iterations, with 10 bootstrap replicates. Layers trimmed to the extent of each species were used; however, all models were projected to the extent defined by all occurrence records. For all models, 25% of the data were set aside for model testing, and the remainder was used for training. Additionally, we disabled extrapolation of extant species geographic projections and allowed missing data. We evaluated each model’s ability to differentiate between suitable and unsuitable areas on the basis of the area under the curve statistic (AUC). Models were further evaluated by calculating jackknife values and response curves.

On the basis of these models, we defined the predicted niche occupancy (PNO) for each variable, as calculated using pno_calc (https://github.com/ryanafolk/pno_calc). PNO profiles represent the distribution of occupancy in niche space for each ecological variable by binned probability distributions. We then calculated a weighted mean for each variable. To investigate fundamental niche in a phylogenetic context, we implemented a phylogenetic PCA (pPCA) with Brownian motion with the function phyl.pca from the R package phytools (Revell, 2012).

To differentiate among modes of speciation, we conducted an age-overlap correlation test, based on the realized niche (Fitzpatrick & Turelli, 2006), using ENMtools (Warren et al., 2010). Niche overlap was quantified as Schoener’s D similarity index, where zero represents no niche similarity and one represents completely identical niches (Schoener, 1968; Warren et al., 2008). The average age overlap at each mode was calculated, and a Monte Carlo permutation test with 1000 simulations was conducted to approximate the P value. To identify the most probable mode of speciation among nodes, we evaluated the intercept and slope when the overlap was plotted against age. Sympatric speciation was chosen as likely if the intercept was found to be greater than 0.5; alternatively, allopatric (or parapatric) speciation was identified as probable if the intercept was less than 0.5. Differentiation of allopatric and parapatric speciation is based on the slope; parapatric speciation is associated with a slope that is near or below zero, whereas allopatric speciation is associated with a slope that is positive (Fitzpatrick & Turelli, 2006). The difference in inferred overlap based on points and range can, therefore, help infer the probable mode of speciation (Cardillo & Warren, 2016). We repeated this analysis, based on both range overlap and point overlap, to confirm our findings.

Using the PNO profiles, we examined phylogenetic niche conservatism by calculating Blomberg’s K statistic and Pagel’s lambda (λ) with the phylosig function in the R package phytools (Revell, 2012). Blomberg’s K assumes that evolution of traits occurs on the basis of Brownian motion (BM), having a random walk process on the tree (Blomberg et al., 2003). If K is equal to zero, characters have no phylogenetic autocorrelation; for K less than one, the traits are less similar than expected under BM; for K greater than or equal to one, the traits are more similar among tips than expected under BM. Pagel’s lambda measures the similarity of covariance among species relative to the correlation expected under BM. For λ equal to zero, the traits have no phylogenetic signal, whereas λ equal to one indicates trait evolution expected under BM (Pagel, 1999). Notably, λ cannot detect phylogenetic signal stronger than BM expectations (Molina-Venegas et al., 2015).

To assess ecological niche shifts, we used the R package hou (Khabbazian et al., 2016), which detects the location and number of trait shifts based on LASSO (Least Absolute Shrinkage and Selection Operator) (Tibshirani, 1996). We did not use our pPCA results to decrease the dimensions of our data, because the top principal component is influenced heavily by traits that occur early in the tree (Uyeda et al., 2015). Instead, we investigated shifts based on a combined matrix of all seven climatic traits and for each trait independently. We estimated shift configuration using phylogenetic-aware information criteria (pBIC) to select the best model, which accounts for both phylogenetic correlation and a large number of possible shift configurations (Khabbazian et al., 2016). We used a random root covariance model with
3 Results

3.1 Phylogenetic relationships
In all, 189 loci were assembled, including 78 nuclear genes and 111 plastid regions (Table S3). After filtering the assembled genes to include only genes assembled for at least 50% of the individuals, our final data set contained 105 genes: 71 nuclear genes and 34 plastid regions. The concatenated alignment with both nuclear and plastid genes contained 31 samples (one to three samples per species; Table S1) and was 77,398 bp in length.

On the basis of the maximum likelihood (ML) inference from the concatenated matrix containing both plastid and nuclear genes, we recovered a generally well-supported tree for Diapensiaceae (21 of 28 nodes had bootstrap support (BS) >90%, six additional nodes had BS >80%; Fig. 1). All genera were monophyletic. Galax was sister to the rest of the family, with 100% BS. Pyxidanthera was sister to a clade of Shortia, Berneuxia, Schizocodon, and Diapensia (100% BS). Diapensia was sister to Schizocodon, Shortia, and Diapensia (100% BS). Diapensia and Schizocodon were sisters (98% BS).

The topology inferred using ML and the nuclear + plastid concatenated matrix differed from inferences based on concatenated analyses of only nuclear or only plastid data (Fig. 2); however, BS for the alternative topologies was low (<75% BS) at conflicting nodes. The overall relationships among genera in the multispecies coalescent tree are similar to those in the ML nuclear + plastid tree, but quartet support was low (Fig. S3). In the coalescent tree, D. wardii was found to be sister to Schizocodon, rather than to other species in Diapensia as in the ML nuclear + plastid tree. In contrast, D. wardii was sister to Shortia galacifolia in the nuclear ML tree. However, given that sequencing of D. wardii only targeted plastid regions, these topologies are based on only one nuclear gene for D. wardii; therefore, these relationships must be viewed with caution.

Within each of the genera with more than one species (Diapensia, Schizocodon, Shortia, and Pyxidanthera), relationships were inconsistent among the coalescent tree, the ML nuclear tree, the ML plastid tree, and the ML nuclear + plastid tree (Figs. 1, 2, S3). Within Diapensia, the ML plastid tree differs from the three other inferences; however, all these relationships had a low BS (<75% BS) and represent soft incongruence. A high BS indicated that D. himalaica and D. purpurea are not reciprocally monophyletic, based on the full nuclear + plastid ML tree and the nuclear ML tree. Sister to D. himalaica and D. purpurea is a clade of D. lapponica and D. obovata; conflicting topologies for these sister relationships were weakly supported (<75% BS) or supported by a low number of concordant genes (Fig. 2). In Schizocodon, Sc. soldanelloides and Sc. ilicifolius are both monophyletic in the nuclear trees, but Sc. ilicifolius is nested within Sc. soldanelloides in the plastid phylogeny. The relationship between S. sinensis and other species of Shortia is unclear; the two samples of S. sinensis did not form a clade. Finally, all inferences suggested that P. barbulata and P. brevifolia are not reciprocally monophyletic.

As we constrained the age of the Diapensiaceae crown group, we could only infer divergence times within Diapensiaceae (Fig. S4). Divergence times of crown groups corresponding to genera within Diapensiaceae spanned the Eocene, the Oligocene, and the early to mid-Miocene (Fig. S4). The age of the clade containing Diapensia, Schizocodon, Shortia, and Berneuxia was estimated to be between 44.8–53.8 Ma. The divergence time of Diapensia and Schizocodon was estimated to be ~27.6 Ma (95% HPD: 21.54–34.14 Ma).

3.2 Biogeographic inferences
Based on AIC values, we found DEC to be the most probable model for our data (Table S4). The crown group was inferred to have originated in the Nearctic with the inferred time of origin of this clade (95% HPD: 64.98–64.99 Ma) (Fig. 3). The probable ancestor of Pyxidanthera, also a North American endemic, was inferred to have occupied the Nearctic region. The ancestor of Shortia, Berneuxia, Schizocodon, and Diapensia likely occurred in the Nearctic region. The probable ancestor of Shortia was predicted to have occurred in the Nearctic, Palearctic, and Indo-Malaysia regions. Shortia, excluding the ancestor of S. galacifolia, likely occurred in the Palearctic and Indo-Malaysia regions. This finding
suggests that S. galacifolia may have separated from the remaining species of Shortia via long-distance dispersal or a vicariance event about 24.8 Mya (95% HPD: 18.87–31.12 Mya). The analysis of regions defined on the basis of the current distribution of species recovered very similar probable ancestors and did not provide additional clarity (Fig. S3).

3.3 Niche evolution
We obtained a total of 5673 raw occurrence records; after cleaning, we retained 2552 occurrence records for analysis. The largest number of occurrence records was obtained for D. lapponica (692) and the fewest for D. wardii (3) (Fig. S1; Table S2). Ecological niche models for all species had AUC scores >0.80; therefore, we considered these models suitable for further analysis. It should be noted that a model was not constructed for D. wardii due to insufficient records.

The two axes of the PCA explained 43.3% and 20.9% of the variance, respectively. For PC1, the annual mean temperature has the largest contribution (20.98%), followed by annual precipitation (19.62%) and precipitation of the coldest quarter (19.41%) (Table S5). For PC2, the largest factor loading is the mean temperature of the wettest quarter (27.13%), followed by mean diurnal range (19.75%) and precipitation seasonality (17.09%) (Table S5). On the basis of the PCA of the realized niche, the current ecological niche space was found to slightly overlap among all but two species indicated in green; the latter two species, D. lapponica and D. obovata, clustered together and occupied broad ranges, including both Nearctic and Palearctic realms (Fig. 4A, represented by green vs. all other colors; Table 1). Similarly, the pPCA based on the abiotic fundamental niche revealed clustering by the biogeographic realm (Fig. 4B; Table 1).
Many climatic variables significantly differed among species in the analysis of the realized niche (Fig. S6; Table S6). In sister group comparisons, we found no significant difference for any climatic variable between \textit{P. barbulata} and \textit{P. brevifolia} or between \textit{D. purpurea} and \textit{D. himalaica}. In contrast, \textit{Sc. soldanelloides} and \textit{Sc. ilicifolius} significantly differed for variables related to precipitation (bio12, bio15, and bio19). In addition, \textit{D. lapponica} and \textit{D. obovata} differed significantly for all climatic variables except the mean diurnal range (bio2).

The mode of speciation was inferred on the basis of the age-overlap correlation test; we found a significant slope and intercept for estimates based on both points and range. Specifically, the overlap at younger nodes was higher than that at older nodes, and the overall slope was negative, indicative of parapatric speciation (Table 2; Fig. S7). We would expect allopatric speciation to have a lower ecological overlap than observed in our phylogeny; however, the calculated \( \lambda \) was greater than one, which may indicate that the rate of evolution of these traits is high at the root of the tree and decreases toward the tips.

We identified five niche shifts along branches with bootstrap support between 63\% and 100\%, based on a combined matrix of all seven climatic variables (Fig. 5). On the basis of separate climatic variables, we found niche shifts for the annual mean temperature (bio1), annual precipitation (bio12), and precipitation of coldest quarter (bio19) (Figs. 5, S8).

4 Discussion

4.1 Phylogeny and divergence times

This study presents a well-supported phylogeny of Diapensiaceae, based on hybrid capture of nuclear genes and plastome sequencing, with a sampling of all 15 species recognized before 2017; two recently described species were not included due to

[Fig. 3. Ancestral range reconstruction from “BioGeoBEARS” for Diapensiaceae where pie charts at nodes indicate the likelihood of originating in a region. Following Rose et al. (2018), regions were categorized as follows: (C) Indo-Malaysia, (D) Nearctic, (E) Neotropics, and (F) Palearctic.]
limited availability of samples. The relationships among genera obtained here are consistent with some of the conclusions of previous studies, notably the subsequent sister relationships of Galax and then Pyxidanthera with the rest of the family (Rönblo & Anderberg, 2002; Higashi et al., 2015). Similar to Higashi et al. (2015), we found that Schizocodon and Shortia are both monophyletic. The general topology we recovered on the basis of combined nuclear and plastid data set agrees with plastome-based inferences for the family (Ye et al., 2020).

As shown in many previous studies, we demonstrate that off-target plastid genes can be readily recovered with Hyb-Seq (e.g., Weitemier et al., 2014; Folk et al., 2015; Carlsen et al., 2018; Herrando-Moraira et al., 2019). On the basis of the ML analysis, we found some discordance between nuclear and plastid trees; however, in these instances, BS for these relationships in the plastid phylogeny was low (<75%). Cytonuclear discordance in these phylogenies could be due to phylogenetic error or to evolutionary processes, which may produce similarly incongruent patterns (e.g., Morales-Briones et al., 2018). For example, many samples had

Table 1 The relative contribution of the first two principal components (PC) for both abiotic realized and fundamental niche

|          | Realized niche | Fundamental niche |
|----------|----------------|-------------------|
|          | PC1  | PC2  | PC1  | PC2  |
| Bio1     | 21   | 12.2 | 14.1 | 18.3 |
| Bio2     | 4.8  | 19   | 11.0 | 7.3  |
| Bio7     | 12.7 | 9    | 19.6 | 12.6 |
| Bio8     | 9.7  | 27.4 | 10.3 | 20.2 |
| Bio12    | 19.6 | 0.4  | 28.2 | 0.1  |
| Bio15    | 12.8 | 17.7 | 1.3  | 21.8 |
| Bio19    | 19.4 | 14.3 | 15.5 | 19.7 |

Variables include the annual mean temperature (bio1), the mean diurnal range (bio2), the temperature annual range (bio7), the mean temperature of the wettest quarter (bio8), annual precipitation (bio12), precipitation seasonality (bio15), and precipitation of coldest quarter (bio19).

Table 2 The age-overlap correlation analysis determined the intercept and slope for ecological overlap, or Schoener’s D, and age (Mya); P-value is indicated in parentheses

| Intercept | Slope |
|-----------|-------|
| Point-based | 0.22 (0.003996*) | -0.004 (0.0099901*) |
| Range-based | 0.395 (0.00599401*) | -0.007 (0.00599401*) |

Intercept supports allopatric or parapatric speciation, whereas the slope supports parapatric speciation. All values are significant (*).
Tests for phylogenetic signal based on Blomberg’s K and Pagel’s lambda (λ)

| Bio1 | 0.612 | 0.249 | 7.85e-5 | -81.843 | -81.843 | 1.00 |
| Bio2 | 1.009 | 0.016* | 1.185 | -59.141 | -60.478 | 0.102 |
| Bio7 | 0.536 | 0.249 | 7.85e-5 | -77.878 | -77.878 | 1.00 |
| Bio8 | 0.668 | 0.147 | 7.85e-5 | -75.781 | -75.781 | 1.00 |
| Bio12 | 0.840 | 0.069 | 1.187 | -99.742 | -104.556 | 0.002* |
| Bio15 | 0.840 | 0.069 | 1.187 | -99.742 | -104.556 | 0.002* |
| Bio19 | 0.850 | 0.036* | 1.188 | -84.742 | -87.669 | 0.015* |

logL is the log-likelihood, whereas logLo is the log-likelihood for lambda equal to zero. Variables include the annual mean temperature (bio1), the mean diurnal range (bio2), the temperature annual range (bio7), the mean temperature of the wettest quarter (bio8), annual precipitation (bio12), precipitation seasonality (bio15), and precipitation of coldest quarter (bio19).

Our molecular dating suggests that the family diversified throughout the Eocene, Oligocene, and into the Miocene. The ancestor of the crown clade most likely originated in the Nearctic. The current geographic distribution is shaped by vicariance with possible dispersal events as well. The disjunction between EA and ENA occurs at two phylogenetic levels in Diapensiaceae. Galax and Pyxidanthera, both from ENA, are subsequent sisters to the rest of the family, most of which occur in EA. Divergence of Pyxidanthera from the clade containing Berneuxia, Schizocodon, Shortia, and Diapensia occurred ~49.9 Mya (95% HPD: 44.82–53.8 Mya). More recently, the divergence between Shortia galacifolia from ENA and the remaining species of Shortia, all from Asia, has occurred ~24.1 Mya (95% HPD: 18.33–30.16 Mya). These disjunctions, therefore, represent pseudocongruence or topological congruence among clades that diverged at different times and could be due to different processes (Cunningham & Collins, 1994; Xiang et al., 1998).

The divergence time between Galax and the rest of the family, as previously estimated by Higashi et al. (2015) in the late Miocene at ~9.5 Mya (95% HPD: 4.9–15.7 Mya), is much too recent and has not been corroborated by subsequent studies. Instead, later studies dated this divergence as much older, at 73.6 Mya (95% HPD: 59.2–88.4 Mya) (Hou et al., 2016) and 59.9 Mya (Rose et al., 2018). The older age estimate of the Diapensiaceae crown (95% HPD: 64.98–64.99 Mya), used here to constrain our divergence time estimates within the family, is corroborated by the age of the fossil Actinocalyx bohrii Friis, which is considered to be closely related to Diapensiaceae. This fossil is dated at 85.8–70.6 Mya and is assigned to the stem of Diapensiaceae and Styracaceae, when considering representative genera Diapensia and Styrax (Friis, 1985; Rose et al., 2018). However, due to the lack of a fossil within the crown clade, the variance around the estimated ages within the family was often large (Fig. S4). We were unable to include additional outgroups due to the absence of available sequence data. The use of additional outgroups within Ericales would allow a new estimation of Diapensiaceae crown node age and may help decrease the divergence time variance at each node.

The realized niche was differentiated between the two broadly distributed species, D. lapponica and D. obovata, and all of the remaining taxa, which are endemic to either EA or ENA. Our analyses suggest that a niche shift accompanied range expansion to this broad distribution of Diapensia in the Northern Hemisphere. Across the entire family, we found a modeled ecological niche to be grouped by a biogeographic realm, indicating that ecological niche may be structured by geography. In contrast, closely related species within the family often occupy the same biogeographic realm, and clustering of these realms in the pPCA indicates that the ecological niche is phylogenetically structured. Among all species of Diapensiaceae, we found that ecological niche has more overlap among species at younger nodes and less overlap at older nodes. For example, Shortia has no ecological niche overlap with other genera, thus suggesting phylogenetic niche divergence between the ancestors of these species and other members of the family.

A low ecological niche overlap was seen between the two monotypic genera, Galax and Berneuxia, and the rest of the family; despite the low overlap, no climatic variables are significantly different for these species, compared with the remaining taxa. Berneuxia thibetica diverged around 30.04 Mya (95% HPD: 23.46–36.27 Mya), and a niche shift for lower precipitation at the coldest quarter was identified for this species.

For the clade of Diapensia lapponica, D. obovata, D. himalaica, and D. purpurea, niche shifts toward lower annual precipitation and lower precipitation of the coldest quarter were detected, relative to their sister group, Schizocodon. Diapensia lapponica and D. obovata diverged from the rest of
Diapensiaceae in the Miocene, around 14.37 Mya (95% HPD: 9.76–18.63 Mya), from an ancestor that occupied the Palearctic region. Despite an overlapping realized niche space, based on the PCA analysis, there is a difference in niches of *D. lapponica* and *D. obovata*, and all bioclimatic variables differ significantly between these sister species, except for mean diurnal range (bio2). Both an overall niche shift to colder and drier habitat and a shift related to a decrease in the annual mean temperature were estimated in the common ancestor of *D. lapponica* and *D. obovata*, relative to the rest of the family. The shift to a colder temperature may have enabled the range expansion of these taxa, because snow and ice have been observed to enable seed dispersal in these species (Day & Scott, 1981). The

Fig. 5. Estimated ecological niche shifts for, A, all environmental variables and, B–D, separate variables (all other variables had zero predicted niche shifts); heatmaps display weighted mean predicted niche occupancy (PNO) values for each species. The bootstrap value is denoted by stars, which indicate predicted shift along a branch. Colored branches indicate niche shifts.
fundamental niche showed phylogenetic niche conservation and high overlap, and parapatric speciation was inferred between *D. lapponica* and *D. obovata*. Despite overlapping broad geographic ranges, a genetic differentiation is still probable among populations of each species of *Diapensia*, and differences in microhabitat requirements may have contributed to speciation. For example, overlapping species of peatmoss (*Sphagnum*) occupying amphi-Beringian and amphi-Atlantic ranges differentiated genetically, likely due to dispersal barriers, and the contemporary distributions of these *Sphagnum* species were shaped by microhabitat variables (Kyrkjeide et al., 2016). The overlapping distributions and similar morphologies of *D. lapponica* and *D. obovata* may have led to some incorrectly identified specimens, so the results of the niche analysis should be interpreted with caution. However, this issue applies more generally to all studies based on natural history collections (Goodwin et al., 2015).

*Diapensia himalaica* and *D. purpurea* were both non-monophyletic, as in a previous study that sampled multiple specimens for each species (e.g., Hou et al., 2016). These two species were monophyletic only in the nuclear concatenated specimens for each species (e.g., Hou et al., 2016). These two monophyletic, as in a previous study that sampled multiple specimens and sampling locations used in our analysis. An overall niche shift was detected along the stem of *Diapensia*, indicating a shift to a colder and wetter habitat. The pPCA revealed that annual precipitation, precipitation seasonality, and precipitation of the coldest quarter (Fig. S6). However, these results are tempered by the possible lack of monophyly for species of *Shortia*, requiring additional investigations into the circumscription and relationships among *Shortia* species. Neither *S. sinensis* nor *S. uniflora* is monophyletic in our analysis, unlike a previous study (Higashi et al., 2015). This result could be due to the collection location of our *S. sinensis* samples (Ha Giang, Vietnam, and Yunnan, China); previous studies included collections only from Yunnan (Higashi et al., 2015). The *S. sinensis* specimen from Vietnam could instead be a separate species, potentially *S. rotata* (Gaddy & Nuraliev, 2017), which was recently recognized as a segregate of *S. sinensis* from Vietnam. In our reduced phylogeny with only one representative per species to maximize sequence coverage per species, we used only the *S. sinensis* specimen from Yunnan. Future investigations into the biogeographical and ecological history of *S. rotata* and *S. sinensis* are needed. In addition, inconsistent recovery of relationships between *S. sinensis* and *S. uniflora* could be due to incomplete lineage sorting or hybridization/introgression.

Within the *Shortia* clade, we found three overall niche shifts for *S. sinensis*, *S. uniflora*, and *S. rotundifolia*; additional shifts were found for increased annual precipitation and decreased precipitation of the coldest quarter for *S. rotundifolia* and *S. sinensis*, respectively. *Shortia rotundifolia* shifted to a slightly warmer and wetter habitat, including increased precipitation during the coldest quarter. *Shortia sinensis* also shifted to a warmer habitat with increased precipitation overall, but with a drier coldest quarter; however, these results for *S. sinensis* were based on only seven occurrence points. Finally, *Shortia uniflora* shifted to a colder and wetter habitat. The pPCA revealed that annual precipitation may separate the fundamental niches of these species. On the basis of proposed poor seed dispersal for members of the family (Ross, 1936; Primack & Wyatt, 1975; Scott, 2004), long-distance dispersal seems unlikely, and vicariance more likely led to the present distributions of species of *Shortia*. *Shortia galacifolia*, endemic to North America, diverged from the Asian sister taxa approximately 24.1 Mya (95% HPD: 18.33–30.16 Mya), consistent with the timing of disjunction in many other taxa (e.g., Wen et al., 2010). The current distribution of the two *Pyxidanthera* species is disjunct within the Atlantic Coastal Plain, with one species occurring in New York and New Jersey and the second species in the south in Virginia, North Carolina, and South Carolina. *Pyxidanthera* was separated into two species, based on differences in leaf morphology (size and pubescence) and habitat (Wells, 1929). Leaf morphological differences between *P. barbulata* and *P. brevifolia* have been related to habitat differences, specifically hydrological differences (Primack & Wyatt, 1975). However, we found that the ecological niches of the two species were not significantly different for any variable, including aspects related to precipitation. Further work is needed to investigate the distinctness of *P. barbulata* and *P. brevifolia*.

On the basis of occurrence records, the species of *Shortia* have no geographic overlap and no ecological niche overlap, and they seem to inhabit different climatic conditions. For example, *S. sinensis* and *S. uniflora* differed significantly in the annual mean temperature, the annual temperature range, annual precipitation, precipitation seasonality, and precipitation of the coldest quarter (Fig. S6). However, these results are tempered by the possible lack of monophyly for species of *Shortia*, requiring additional investigations into the circumscription and relationships among *Shortia* species. Neither *S. sinensis* nor *S. uniflora* is monophyletic in our analysis, unlike a previous study (Higashi et al., 2015). This result could be due to the collection location of our *S. sinensis* samples (Ha Giang, Vietnam, and Yunnan, China); previous studies included collections only from Yunnan (Higashi et al., 2015). The *S. sinensis* specimen from Vietnam could instead be a separate species, potentially *S. rotata* (Gaddy & Nuraliev, 2017), which was recently recognized as a segregate of *S. sinensis* from Vietnam. In our reduced phylogeny with only one representative per species to maximize sequence coverage per species, we used only the *S. sinensis* specimen from Yunnan. Future investigations into the biogeographical and ecological history of *S. rotata* and *S. sinensis* are needed. In addition, inconsistent recovery of relationships between *S. sinensis* and *S. uniflora* could be due to incomplete lineage sorting or hybridization/introgression.
et al., 2010). For example, divergence between EA and ENA relatives was dated to the late Oligocene and early Miocene in Vitaceae (Nie et al., 2010), Saxifragaceae (Deng et al., 2015), and Phrymaceae (Nie et al., 2006).

5 Conclusions

Overall, we provided clarity regarding phylogenetic relationships, biogeography, and ecological niche evolution in Diapensiaceae. A phylogenetic understanding of Diapensiaceae could be improved by increased sampling within species, as we sampled only one to three individuals per species. The resolution of questions of species monophyly may require greater sampling. Additional molecular data could also improve the phylogenetic inference of Diapensiaceae. We did not include subspecies within the family due to specimen limitations. Likewise, two recently described species, Shortia rotata (Gaddy & Nuraliev, 2017) and S. brevistyla (Gaddy et al., 2019), segregates of S. sinensis and S. galacifolia, respectively, were not included because (i) we were unable to obtain specimens for S. rotata, and (ii) S. brevistyla had not been recognized when we collected samples for this study. In addition, herbarium specimens of neither S. rotata nor S. brevistyla have been annotated as distinct from S. sinensis and S. galacifolia, respectively; the digitized herbarium specimen records for S. sinensis and S. galacifolia may, thus, represent lumped entities. Therefore, S. rotata and S. brevistyla could not be included as separate species in our analyses.

Many questions remain unanswered regarding the evolution of Diapensiaceae. Scott & Day (1983) noted the low chromosome number for the family—most members of Diapensiaceae have $n = 6$ ($2n = 12$), with Galax urceolata having $2n = 12$, 18, or 24 (Reynolds, 1968)—however, chromosome numbers across the complete range of most species in Diapensiaceae are unknown, and chromosome evolution within the family has not been thoroughly investigated. Scott & Day (1983) also suggested that further attention should be paid to the factors governing range size, given that dispersal appears to be limited in the majority of taxa analyzed. Although we identified that temperature shifts may have allowed D. iaponica and D. obovata to expand their ranges, much is still unknown about the persistence of the taxa through climate oscillations. Also, the effect of climate change on the family should be investigated. For example, climate change is predicted to greatly decrease the range of G. urceolata (Gaynor et al., 2018), and similar analyses and projections would be useful for other species in the family. We hypothesize that other species of Diapensiaceae, given their general dependence on cool temperatures, may be similarly impacted by a rapidly changing climate.

The current distributions of species of Diapensiaceae are shaped by both climatic and biogeographic factors. Disjunctions between clades recognized as genera are most likely due to ancient vicariance events. On the basis of our inferences, the disjunction between ENA and EA in Diapensiaceae can be traced to two vicariance events—(i) Pyxidanthera versus the clade of Berneuxia, Shortia, Schizocodon, and Diapensia at 49.9 Mya (95% HPD: 44.82–53.8 Mya) and (ii) a split between Shortia galacifolia and all other Shortia species around 24.1 Mya (95% HPD: 18.33–30.16 Mya). These findings clearly provide further examples of pseudocongruence across the well-known ENA–EA disjunction. Although all species of the family occupy generally similar habitats, as seen by analysis of specific ecological variables as well as ecological niche models, niche shifts have accompanied some divergences, such as the origin of Schizocodon and, more recently, between species of Schizocodon and among EA species of Shortia. Overall, the diversification in Diapensiaceae appears to have been shaped by both large-scale biogeographic factors, such as vicariance, and divergence in ecological niche among closely related species.

Acknowledgements

The authors thank Ryan A. Folk, Johanna R. Jantzen, Rhett M. Rautsaw, Andrew Mason, Eric W. Goolsby, Anthony E. Melton, Grant T. Godden, Andre A. Naranjo, and Jacob B. Landis for their guidance regarding data analysis. They also acknowledge Matthew A. Gitzendanner, Kent D. Perkins, and William M. Whitten for their assistance in the lab and/or the FLAS herbarium. The greenhouse sample of Galax urceolata was provided by Xiang Liu and Wenbin Zhou, and the collection was approved by Mark John (Hemlock Bluffs Natural Preserve). Occurrence data from the Chinese Virtual Herbarium were cleaned by Haihua Hu. Funding was provided by iDigBio (NSF grant number DBI-1547229), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000), the Large-scale Scientific Facilities of the Chinese Academy of Sciences (Grant No: 2017-LSFGBOWS-01), and NSF Dimensions of Biodiversity grant DEB-1442280.

References

Andermann T, Cano Á, Zizka A, Bacon C, Antonelli A. 2018. SECAPR—A bioinformatics pipeline for the rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments. PeerJ 6: e5175.

APG IV, Chase MW, Christenhusz MJM, Fay MF, Byng JW, Judd WS, Soltis DE, Mabberley DJ, Sennikov AN, Soltis PS, Stevens PF. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society 181: 1–20.

Baldwin J, Jr. 1939. Chromosomes of the Diapensiaceae: A cytological approach to a phylogenetic problem. Journal of Heredity 30: 169–171.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prijibelski AD. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19: 455–477.

Barringer BC, Galloway LF. 2017. The reproductive ecology of diploid and tetraploid Galax urceolata. The American Midland Naturalist 177: 299–308.

Bivand R, Rundel P, Pebesma E, Stuetz R, Hufthammer KO, Giraudoux P, Davis M, Santilli S, Bivand MR. 2019. Package ‘rgdal’ [online]. Available from https://CRAN.R-project.org/package=rgdal.
Hijmans RJ, van Etten J, Mattiuizi M, Sumner M, Greenberg J, Lamignueiro O, Bevan A, Racine E, Shortridge A. 2013. Raster package in R. Version 3.0-12 [online]. Available from https://CRAN.R-project.org/package=raster.

Hong D-Y. 1993. Eastern Asian–North American disjunctions and their biological significance. Cathaya 5: 1–39.

Hou Y, Bjorå CS, Ikeda H, Brochmann C, Popp M. 2016. From the north into the Himalayan–Hengduan Mountains: Fossil-calibrated phylogenetic and biogeographical inference in the arctic-alpine genus Diapensia (Diapensiaceae). Journal of Biogeography 43: 1502–1513.

Hou Y, Nowak MD, Mirré V, Bjorå CS, Brochmann C, Popp M. 2015. Thousands of RAD-seq loci fully resolve the phylogeny of the highly disjunct arctic-alpine genus Diapensia (Diapensiaceae). PLoS One 10: e0140175.

Hultén E. 1958. The amphi-atlantic plants and their phytogeographical connections. Kungliga Svenska Vetenskapsakademiens Handlingar, Fjärde Serien, 7: 1–340.

Hultén E, Fries M. 1986. Atlas of North European vascular plants north of the Tropic of Cancer. Oberreifenberg: Koeltz Scientific.

Johnson LA, Soltis DE. 1998. Assessing congruence: Empirical examples from molecular data. In: Soltis DE, Soltis PS, Doyle JJ eds. Molecular systematics of plants II. Boston, MA: Springer. 287–348.

Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJ, Wickett NJ. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. Applications in Plant Sciences 4: 1600016.

Johnson MG, Pokorny L, Dodsworth S, Botigué LR, Cowan RS, Deavult A, Eiserhardt WL, Epitaphe N, Forest F, Kim JT. 2018. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. Systematic Biology 68: 594–606.

Junier T, Zdanov EM. 2010. The Newick utilities: High-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26: 1669–1670.

Katoh K, Misawa K, Kuma Y, Miyata T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30: 3059–3066.

Khabbazian M, Kriebel R, Rohe K, Ané C. 2016. Fast and accurate phylogenetic inference in large genomic alignments. Proceedings of the 21st International Conference on Machine learning, ACM. 83.

Kyrkjeeide MO, Hassel K, Flatberg KI, Shaw AJ, Brochmann C, Stenåen HK. 2016. Long distance dispersal and barriers shape genetic structure of peatmosses (Sphagnum) across the Northern Hemisphere. Journal of Biogeography 43: 1215–1226.

Landis MJ, Matzke NJ, Moore BR, Huelsenebeck JP. 2013. Bayesian analysis of biogeography when the number of areas is large. Systematic Biology 62: 789–804.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. Partition Finder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29: 1695–1701.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25: 1754–1760.

Li HL. 1952. Floristic relationships Eastern Asia and Eastern North America. Transactions of the American Philosophical Society 42: 371–429.

MacKenzie CM, Murray G, Primack R, Weihrauch D. 2017. Lessons from citizen science: Assessing volunteer-collected plant phenology data with Mountain Watch. Biological Conservation 208: 121–126.

Malychev L. 1997. Flora of Siberia, Volume 11: Pyrolaceae–Lamiaceae (Labiatae). Enfield: Science Publishers.

Manos PS, Meireles JE. 2015. Biogeographic analysis of the woody plants of the Southern Appalachians: Implications for the origin of a regional flora. American Journal of Botany 102: 780–804.

Matzke NJ. 2013. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. R Package, Version 0.2. 1. 2013.

Molina-Venegas R, Aparicio A, Slingsby JA, Lavergne S, Arroyo J. 2015. Investigating the evolution of the Mediterranean biodiversity hotspot: Deep phylogenetic signal in the distribution of eucaulis across elevational belts. Journal of Biogeography 42: 507–518.

Morales-Briones DF, Liston A, Tank DC. 2018. Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus Lachemilla (Rosaceae). New Phytologist 218: 1668–1684.

Nie ZL, Sun H, Beardsley PM, Olmstead RG, Wen J. 2006. Evolution of biogeographic disjunction between eastern Asia and eastern North America in Phryma (Phrymaceae). American Journal of Botany 93: 1343–1356.

Nie ZL, Sun H, Chen ZD, Meng Y, Manchester SR, Wen J. 2010. Molecular phylogeny and biogeographic diversification of Parthenocissus (Vitaceae) disjunct between Asia and North America. American Journal of Botany 97: 1342–1353.

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV, Underwood EC, D’amico JA, Itoua I, Strand HE, Morrison JC. 2001. Terrestrial ecoregions of the world: A new map of life on earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience 51: 933–938.

Pagel M. 1999. Inferring the historical patterns of biological evolution. Nature 401: 877–884.

Peterson AT, Soberón J, Pearson RG, Anderson RP, Martínez-Meyer E, Nakamura M, Araújo MB. 2011. Ecological niches and geographic distributions (MPB-49). Princeton: University Press.

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

Phillips SJ, Dudík M, Schapire RE. 2004. A maximum entropy approach to species distribution modeling. Proceedings of the 21st International Conference on Machine learning, ACM. 83.

Portik DM, Wiens JJ. 2019. SuperCRUNCH: A toolkit for creating and manipulating supermatrices and other large phylogenetic data sets. bioRxiv.

Primack RB, Wyatt R. 1975. Variation and taxonomy of Pyxidanthera (Diapensiaceae). Brittonia 27: 115–118.

Pyron R, Alexander R, Costa GC, Patten MA, Burbink FT. 2015. Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. Biological Reviews 90: 1248–1262.

Qian H, Ricklefs RE. 1999. A comparison of the taxonomic richness of woodland plants of the Southern Appalachians: Implications for the origin of the Appalachian genus Pyxidanthera (Diapensiaceae). Brittonia 51: 250–259.

Underwood EC, D’amico JA, Itoua I, Strand HE, Morrison JC.

R Development Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3–900051–07–0 [online]. Available from http://www.R-project.org.

R Development Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3–900051–07–0 [online]. Available from http://www.R-project.org.

Ree RH, Sanmartín I. 2018. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its
comparison with DEC via model selection. *Journal of Biogeogra-
phy* 45: 741–749.

Ree RH, Smith SA. 2008. Maximum likelihood inference of
geographic range evolution by dispersal, local extinction,
and cladogenesis. *Systematic Biology* 57: 4–14.

Revell LJ. 2012. phytools: An R package for phylogenetic comparative
biology (and other things). *Methods in Ecology and Evolution* 3:
217–223.

Reynolds JD. 1968. Morpholohival Studies in Diapensiaceae. I.
Chromosome number and microspore development in *Pyxid-
thera brevifolia* Wells. *Torrey Botanical Society* 95: 653–656.

Römbke K, Anderberg AA. 2002. Phylogeny of Diapensiaceae based
on molecular data and morphology. *Systematic Botany* 27:
383–395.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic
inference under mixed models. *Bioinformatics* 19: 1572–1574.

Rose JP, Kleist TJ, Löfstrand SD, Drew BT, Schönenberger J, Sytsma
KJ. 2018. Phylogeny, historical biogeography, and diversification
of angiosperm order Ericales suggest ancient Neotropical and
East Asian connections. *Molecular Phylogenetics and Evolution*
122: 59–79.

Ross M. 1936. Seed reproduction of *Shortia galacifolia*. *Journal of
the New York Botanical Garden* 37: 208–211.

Sanderson MJ. 2002. Estimating absolute rates of molecular
evolution and divergence times: A penalized likelihood
approach. *Molecular Biology and Evolution* 19: 101–109.

Sanmartin I, Enghoff H, Ronquist F. 2001. Patterns of animal
dispersal, vicariance and diversification in the Holarctic.
*Biological Journal of the Linnean Society* 73: 345–390.

Schoener TW. 1968. The Anolis lizards of Bimini: Resource
partitioning in a complex fauna. *Ecology* 49: 704–726.

Schönenberger J, Anderberg AA, Sytsma KJ. 2005. Molecular
phylogenetics and patterns of floral evolution in the Ericales.
*International Journal of Plant Sciences* 166: 265–288.

Scott PI. 2004. *Diapensiaceae.* In: Kubitzki K ed. *Flowering plants.
Dicotyledons. The families and genera of vascular plants.* Berlin:
Springer. 117–121.

Scott PJ, Day RT. 1983. *Diapensiaceae: A review of the taxonomy.
Taxon* 32: 417–423.

Seelanan T, Schnabel A, Wendel JF. 1997. Congruence and consensus
in the cotton tribe. *Systematic Botany* 22: 259–290.

Setoguchi H, Obha H. 1995. Phylogenetic relationships in *Crosso-
stylosis* (Rhizophoraceae) inferred from restriction site variation
of chloroplast DNA. *Journal of Plant Research* 108: 87–92.

Shen W, Le S, Li Y, Hu F. 2016. SeqKit: A cross-platform and ultrafast
toolkit for FASTA/Q file manipulation. *PLoS One* 11: e0163962.

Shepherd LD, McLay TG. 2011. Two micro-scale protocols for the
isolation of DNA from polysaccharide-rich plant tissue. *Journal of
Plant Research* 124: 311–314.

Smith SA, Moore MJ, Brown JW, Yang Y. 2015. Analysis of
phylogenomic datasets reveals conflict, concordance, and gene
duplications with examples from animals and plants. *BMC Evolutionary
Biology* 15: 150.

Smith SA, O’Meara BC. 2012. *treePL: Divergence time estimation
using penalized likelihood for large phylogenies.* *Bioinformatics* 28:
2689–2690.

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

Stull GW, Soltis PS, Soltis DE, Gitzendanner MA, Smith SA. 2020.
Nuclear phylogenomic analyses support novel relationships
among major asterid lineages. *American Journal of Botany* 107:
790–805.

Tibshirani R. 1996. Regression shrinkage and selection via the lasso.
*Journal of the Royal Statistical Society: Series B (Methodological)*
58: 267–288.

Tiffney BH. 1985a. Perspectives on the origin of the floristic similarity
between eastern Asia and eastern North America. *Journal of
the Arnold Arboretum* 66: 73–94.

Tiffney BH. 1985b. The Eocene North Atlantic land bridge: Its
importance in Tertiary and modern phytogeography of the
Northern Hemisphere. *Journal of the Arnold Arboretum* 66:
243–273.

Tiffney BH, Manchester SR. 2001. The use of geological and
paleontological evidence in evaluating plant phylogeographic
hypotheses in the Northern Hemisphere Tertiary. *International
Journal of Plant Sciences* 162: 53–517.

Uyeda JC, Caetano DS, Pennell MW. 2015. Comparative analysis of
principal components may be misleading. *Systematic Biology* 64:
677–689.

Walker JF, Walker-Hale N, Vargas OM, Larson DA, Stull GW. 2019.
Characterizing gene tree conflict in plastome-inferred phyloge-
nies. *PeerJ* 7: e7747.

Wall WA, Douglas NA, Xiang QY, Hoffmann WA, Wentworth TR,
Hohmann MG. 2010. Evidence for range stasis during the latter
Pleistocene for the Atlantic Coastal Plain endemic genus,
*Pyxidanthera Michaux.* *Molecular Ecology* 19: 4302–4314.

Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalen-
cy versus conservatism: Quantitative approaches to niche evolu-
tion. *Evolution: International Journal of Organic Evolution* 62:
2868–2883.

Warren DL, Glor RE, Turelli M. 2010. *ENMTools: A toolbox for
comparative studies of environmental niche models.* *Ecography* 33:
607–611.

Weihe A, Apitz J, Pohlheim F, Salinas-Hartwig A, Börner T. 2009.
Biparental inheritance of plastidial and mitochondrial DNA and
hybrid variegation in *Pelargonium.* *Molecular Genetics and
Genomics* 282: 587–593.

Weitemier K, Straub SC, Cronn RC, Fishbein M, Schmickl R,
Warren DL, Glor RE, Turelli M. 2014. *HybSeq: Combining target
enrichment and genome skimming for plant phylogenomics.*
*Applications in Plant Sciences* 2: 1400042.

Wells B. 1929. A new pyxie from North Carolina. *Journal of the Elisha
Mitchell Scientific Society* 44: 238–239.

Wen J. 1999. Evolution of eastern Asian and eastern North American
disjunct distributions in flowering plants. *Annual Review of
Ecology and Systematics* 30: 421–455.

Wen J, Ickert-Bond S, Nie Z-L, Li R. Timing and modes of evolution of
eastern Asian-North American biogeographic disjunctions in seed
plants. Darwin’s heritage today. *Proceedings of the Darwin 2010
Beijing International Conference,* 2010. Beijing: Higher Education
Press. 252–269.

Wen J, Nie Z-L, Ickert-Bond SM. 2016. Intercontinental disjunctions
between eastern Asia and western North America in vascular
plants highlight the biogeographic importance of the Bering land
bridge from late Cretaceous to Neogene. *Journal of Systematics
and Evolution* 54: 469–490.

Wiens JJ, Graham CH. 2005. Niche conservatism: Integrating
evolution, ecology, and conservation biology. *Annual Review of
Ecology, Evolution,* and *Systematics* 36: 519–539.
Wolfe JA. 1975. Some aspects of plant geography of the Northern Hemisphere during the late Cretaceous and Tertiary. *Annals of the Missouri Botanical Garden* 62: 264–279.

Xiang Q-Y, Soltis DE. 2001. Dispersal-vicariance analyses of intercontinental disjuncts: Historical biogeographical implications for angiosperms in the Northern Hemisphere. *International Journal of Plant Sciences* 162: 529–539.

Xiang Q-Y, Soltis DE, Soltis PS. 1998. The eastern Asian and eastern and western North American floristic disjunction: Congruent phylogenetic patterns in seven diverse genera. *Molecular Phylogenetics and Evolution* 10: 178–190.

Ye J-F, Niu Y-T, Feng Y-L, Liu B, Hai L-S, Wen J, Chen Z-D. 2020. Taxonomy and biogeography of *Diapensia* (Diapensiaceae) based on chloroplast genome data. *Journal of Systematics and Evolution* 58: 696–709.

Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19. doi: 10.1186/s12859-018-2129-y.

Zicka A, Silvestro D, Andermann T, Azevedo J, Duarte Ritter C, Edler D, Farooq H, Herdean A, Ariza M, Scharn R. 2019. CoordinatedCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* 10: 744–751.

### Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12646/supplinfo:

**Table S1.** Herbarium specimens used in this study as sources of DNA. Specimens were obtained from herbaria including the California Academy of Sciences (CAS), Carnegie Museum of Natural History (CM), Florida Museum of Natural History (FLAS), Idaho State University (IDS), Missouri Botanical Garden (MO), and University of North Carolina at Chapel Hill (NCU). One living specimen was grown in the Biology greenhouse at the University of Florida. Additional sequence data were obtained from Drs. Lianning Gao and Chaonan Fu (Kunming Institute of Botany) (Vouchers: GLM; Sequencing referred to as Plastid). An outgroup sample (Cyrilla racemiflora) was obtained from the National Center for Biotechnology Information Short Read Achieve (Sequencing: SRA) (Voucher: SRR).

**Table S2.** Number of occurrence records obtained for each species.

**Table S3.** Alignment information for each data grouping. Filter relates to alignment filters, 50 corresponds to filtered genes to include genes found in only 50% of the individuals. G is gene type: A indicates all, N indicates nuclear, P indicates plastid. N is the number of samples, while n is the number of genes. %GC is percentage GC content, %PI is percentage pairwise identity among sequences, %IS indicates the percentage of identical sites.

**Table S4.** Model evaluation statistics for ancestral area reconstruction in “BioGeoBEARS”. For each model, log-likelihoods (LnL), degrees of freedom of the null (Df null), and the Akaike Information Criterion (AIC) are indicated.

**Table S5.** Relative contribution of the climatic layers annual mean temperature (bio1), mean diurnal range (bio2), temperature annual range (bio7), mean temperature of wettest quarter (bio8), annual precipitation (bio12), precipitation seasonality (bio15), and precipitation of the coldest quarter (bio19).

**Table S6.** An ANOVA of the environmental variables among species; degrees of freedom (df), F-statistic, and p-value.

**Fig. S1.** Distribution of occurrence records for each species of Diapensiaceae. Color of points indicates realm (see Figure 2) occupied: red (Indo-Malaysia), blue (Nearctic), golden (Palearctic), orange (Indo-Malaysia & Palearctic), and green (Nearctic & Palearctic).

**Fig. S2.** Approximate location of samples used as sources of DNA.

**Fig. S3.** Multi-species coalescent inference based on subset of gene trees for only nuclear genes. Quartet support indicated at nodes. Pie charts indicate the percentage of gene trees that are concordant (blue), top alternative bipartition (green), any other alternative bipartition (red), or uninformative (grey) at that node. Numbers above and below branches, respectively, indicate the number of concordant gene trees and the number of conflicting trees.

**Fig. S4.** Time-calibrated phylogenetic tree, inferred with penalized likelihood, for Diapensiaceae. Blue bars represent the posterior means, or 95% credibility interval, of the node age estimate. Time increments indicated are in millions of years.

**Fig. S5.** Ancestral range reconstruction from ‘BioGeoBEARS’ for Diapensiaceae where pie charts at nodes indicate likelihood of originating in a region. Regions were categorized as follows: (A) Western Asia, (B) Eastern Asia, (C) Scandinavia, (D) Eastern North America, (E) Alaska and Canada, (F) Central and South America. DEC was found to be the most probable model with an AIC of 61.76 (AIC: DIVALIKE, 65.34; BAYAREALIKE, 68.42).

**Fig. S6.** For each bioclimatic variable, we analyzed niche among included species using analysis of variance (ANOVA) followed by post hoc tests using the Tukey Honest Significant Differences (HSD.test). Significantly different (alpha = 0.05) groups are indicated by color and letter for each variable.

**Fig. S7.** Age-overlap correlation test. (A) Overlap, or Schoener’s D similarity index, at each node; top value is point-overlap, bottom value is range-overlap; (B) Point-based age-overlap correlation; (C) Range-based age-overlap correlation.

**Fig. S8.** Estimated niche shift for each climatic variable. Bootstrap support indicated for each estimated shift.