Molecular Phylogenetics and Evolution 95 (2016) 183–195

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Evolution of geographical place and niche space: Patterns of diversification in the North American sedge (Cyperaceae) flora ☆

Daniel Spalink [a,*], Bryan T. Drew [b], Matthew C. Pace [a,1], John G. Zaborsky [a], Pan Li [c], Kenneth M. Cameron [a], Thomas J. Givnish [a], Kenneth J. Sytsma [a]

[a] Department of Botany, University of Wisconsin-Madison, 430 Lincoln Drive, Madison, WI 53706, USA
[b] Department of Biology, University of Nebraska at Kearny, 905 West 25th Street, Kearny, NE 68849, USA
[c] College of Life Sciences, Zhejiang University, #388 Yuhangtang Rd., Xihu District, Hangzhou 310058, Zhejiang, China

ARTICLE INFO

Article history:
Received 28 February 2015
Revised 10 September 2015
Accepted 17 September 2015
Available online 17 December 2015

Keywords:
Biogeography
Museum collections
Phylogenetic niche conservatism
Phylogenetic signal
Rates of evolution
Range and niche evolution

ABSTRACT

The role of geography and ecology in speciation are often discussed in the context of phylogenetic niche conservatism (PNC), the propensity of lineages to retain ancestral niche related traits. However, a recent paradigm shift focuses instead on measuring divergence of these traits in conjunction with patterns of speciation. Under this framework, we analyzed the diversification of North America's third most diverse family, Cyperaceae ("sedges"), using a modified Parsimony Analysis of Endemicity approach to identify floristic regions and ordination statistics to quantify species distribution in a continuous manner. Utilizing over 200,000 georeferenced specimens, we characterized the geographical distribution and climatic and edaphic niche space occupied by each species. We constructed a supermatrix phylogeny of the North American sedge flora, aided in part by the sequencing of all sedges of Wisconsin, and employed a multifaceted approach to assess the role of geographical and ecological divergence on lineage diversification. In addition to measuring phylogenetic signal for these traits, we also measured pairwise phylogenetic distance of species within floristic regions, calculated rates of speciation, and tested for correlations of speciation rate to tempo of geographical and ecological evolution. Our analyses consistently show that evolutionarily related species tend to be geographically unrelated. Rates of geographical and ecological diversification are closely linked to tempo of speciation, and exploration of geographical place coincides with divergence in ecological niche space. We highlight the benefits of treating geography in a continuous manner, and stress the importance of employing a diverse suite of analytical approaches in testing hypotheses regarding the evolution of range and niche.

1. Introduction

The testing of explicit evolutionary hypotheses by integrating molecular, morphological, genetic, geographical, and ecological data has advanced our understanding of the process, timing, location, tempo, and correlates of diversification across many lineages (e.g., Slater et al., 2010; Rabosky et al., 2013; Givnish et al., 2014, 2015). Museum collections (Funk and Richardson, 2002; Pyke and Ehrlich, 2010; Lavoie, 2013; Rocha et al., 2014) as well as repositories of molecular (e.g., NCBI; Geer et al., 2010), geographical (e.g., GBIF; www.gbif.org), and climatic (e.g., BioClim; Hijmans et al., 2005) data have made the acquisition of large amounts of information for this purpose possible, corresponding with a renewed interest in topics such as phylogenetic niche conservatism (PNC) and the role of geography in lineage diversification (Losos, 2008; Wiens et al., 2010; Araya et al., 2012; Crisp and Cook, 2012; Loera et al., 2012). Phylogenetic niche conservatism, the propensity of lineages to retain ancestral niche-related traits, is a controversial topic (Ackerly, 2009; Crisp and Cook, 2012). Whether PNC is an evolutionary process itself or instead a pattern that can be observed as a result of other evolutionary forces (Crisp and Cook, 2012), whether phylogenetic signal of niche traits and deviation from a null Brownian model of trait evolution are evidence of PNC (Ackerly, 2009; Boucher et al., 2014), and whether climatic niches are even appropriate to analyze in a phylogenetic context (Grandcolas et al., 2011), remain topics of debate in this vibrant and important field of inquiry.

In an attempt to overcome some of these challenges, a recent shift in paradigm instead focuses on measuring patterns of niche divergence in relation to lineage diversification and speciation

☆ This paper was edited by the Associate Editor Stefanie M. Ickert-Bond.

* Corresponding author.
E-mail address: dspalink@wisc.edu (D. Spalink).
[1] Present address: The New York Botanical Garden, 2900 Southern Boulevard, Bronx, NY 10458, USA.

http://dx.doi.org/10.1016/j.ympev.2015.09.028
1055-7903/

(Hua and Wiens, 2013; Boucher et al., 2014; Donoghue and Edwards, 2014; Joly et al., 2014a). Under this framework, the role of biogeography becomes increasingly important, particularly for lineages with easily dispersed seeds where migration would facilitate exploration of niche space, serve as a means of reproductive isolation, and ultimately result in lineage splitting. The evolution of species' geographical distributions could either occur in conjunction with niche evolution, such that species disperse and adapt to new places that are ecologically dissimilar from their ancestral areas, or independently of niche divergence, such that species only disperse to and survive in ecologically similar areas. In either scenario, if geography historically served as a means of reproductive isolation, we would expect to see closely related species occupying geographically disparate regions and distantly related species converging in geographically similar regions (i.e., phylogenetic overdispersion; Villalobos et al., 2013). We would also expect to see rates of speciation reflected in rates of range evolution, with rapidly diversifying lineages exploring geographic locales at a faster rate than more slowly diversifying lineages (Ackerly, 2009). However, if species' distributions are independent of past speciation events, we would expect to see closely related species inhabiting congruent, overlapping, or contiguous distributions, such that range *per se* would exhibit a strong phylogenetic signal (i.e., phylogenetic clustering; Webb et al., 2002). Furthermore, we would expect to find no correlation between speciation rate and rate of geographical evolution in these lineages.

Incorporating biogeography into tests of PNC or niche evolution is challenging, as multidimensional geographic space is difficult to characterize in a continuous manner suitable for rigorous phylogenetic comparative analyses. For example, the presence or absence of a taxon in a predefined region, such as a continent, biome, or floristic region, is often utilized in a phylogenetic context for historical biogeographical reconstructions (e.g., Givnish et al., 2014). However, in reality there are few instances in which two species exhibit perfectly congruent distributions, especially as would be captured by this coarse characterization of geographic space. Methods that do treat biogeography in a continuous fashion often include convex-hull (Cornwell et al., 2006) or species distribution modeling (e.g., Joly et al., 2014a) approaches. These approaches can be useful, for example to measure the extent to which species' ranges overlap, but the multidimensionality of such data are difficult to analyze phylogenetically. Although geography is a fundamental component of evolution, it remains a significant challenge to characterize species distribution in a manner sufficient to test a wide variety of evolutionary hypotheses.

Here, we propose two new approaches to utilize geographical data within a phylogenetic context. In the first approach, we divide the geographic region of interest into quadrats and score the presence and absence of species within each quadrat to form a binary matrix. We then conduct correspondence analysis on this matrix to create continuous variables that represent a portion of the variance associated with distribution. We use these variables in conjunction with climatic and edaphic traits to conduct a suite of analyses, including testing for phylogenetic signal (Blomberg et al., 2003) and measuring rates of speciation and trait evolution (Rabosky et al., 2013). In the second approach, we use the same quadrat matrix but instead subject it to phylogenetic analysis and identify "clades" of species that are geographically similar (i.e., floristic regions). We then measure levels of phylogenetic clustering or overdispersion within each of the floristic regions.

We utilize both approaches to determine the role of geography and niche in the evolution of North America's third most diverse family, the Cyperaceae Juss. ("sedges"). The North American sedge flora is comprised of ~843 species in ~24 genera (Ball et al., 2003), generally characterized by low statured, perennial, small-seeded, and wind pollinated and dispersed herbs. Sedges are found, and often dominant, in wetlands of all forms but also in savannas, tundra, rocky outcrops, and forest understories. They are distributed from coast to coast and from seashore to mountaintop, exhibit a remarkable diversity of tolerance to both climate and soil types, and include acidophytes, calciphytes, and halophytes. While some species are widespread and weedy, many exhibit fidelity to particular ecological conditions and are narrowly restricted to small ranges, limited to a single state or only a handful of populations. The extent to which these patterns are the product of sympatric speciation, resulting from either intrinsic or extrinsic barriers to reproduction and fine-scale niche partitioning, or allopatric speciation, resulting from geographical isolation and subsequent genetic and ecological divergence, remains largely unexplored. This former scenario has received some attention, as rates of evolution in the genus *Carex* L., which comprises over half of the North American sedge flora, are associated with chromosome evolution (Escudero et al., 2012). Nevertheless, geographical causes of reproductive isolation cannot be entirely ruled out, as several recent studies have demonstrated the influence of dispersal on the diversification of the sedges on a global scale (Escudero et al., 2009; Viljoen et al., 2013; Villaverde et al., 2015).

To explore these issues, we present a supermatrix phylogeny of the North American Cyperaceae, enabled in part by *de novo* sequencing of all sedges of Wisconsin. We utilize this phylogeny to address the role of geography in the evolution of the species rich and ecologically important North American sedge flora. First, we test the hypothesis that the diversification of sedges was facilitated, at least in part, by range evolution. Second, we assess the extent to which speciation is associated with the evolution of niche, here characterized by both climatic and edaphic variables. Third, we determine if the exploration of geographic place (i.e., migration) coincides with adaptation to novel ecological space.

2. Methods

2.1. Taxonomic and molecular sampling

Through a combination of *de novo* sequencing and data mining from GenBank, we constructed a phylogeny of the North American sedge flora using a modified supermatrix approach (de Queiroz and Gatesy, 2007; Hinchliff and Roalson, 2013). As part of an ongoing effort to produce sequences for the entire vascular flora of Wisconsin, we sequenced four chloroplast DNA gene regions for all Wisconsin sedges (*matK, ndhF, rbcL,* and *trnL-F*), which is comprised of 235 species and represents more than a quarter of all North American sedges. We then augmented this core data set by further sequencing for members of tribe Scirpeae the single copy nuclear gene *phyC*, the *ITS* region of nuclear rDNA, and the cpDNA *rps16* gene. We used primers B49317 and A50272 from Taberlet et al. (1991) for *trnL-F*, *rbcLaF* and *rbcLaR* from Kress and Erickson (2007) for *rbcL*, *ndhF* primers from Gilmour et al. (2013), rpsF and rps2R (Paton et al., 2004) for *rps16*, Leu1 (Andreasen et al., 1999) and ITS4 (White et al., 1990) for *ITS*, and phyC 2F and 2R (Jabaily and Sytsma, 2010) for *phyC*. We designed new primers for *matK*. These include matK5Bf (5′-CGT ACT GTA CTT TTA TGT TTA C-3′) and matK6Ar (5′-ATC CTG TCC ATT TTG AAA TCT TAG-3′). DNA extraction, PCR, and Sanger sequencing followed protocols as outlined in Drew and Sytsma (2012).

We subsequently mined GenBank for sequence data for all North American sedges and *Hypolytrum* Pers., which we used as an outgroup. In all molecular phylogenetic studies of Cyperaceae to date, the genus *Hypolytrum* is placed in the subfamily Mapanioideae, which does not contain any North American taxa and is sister to subfamily Cyperoideae, which contains all North American species (Muasya et al., 2009; Escudero and Hipp, 2013; Hinchliff

and Roalson, 2013). We limited the search to gene regions for which data were available for a minimum of 15 species. We generated phylogenetic trees for each region individually to identify mislabeled sequences, which were eliminated from downstream analyses. When two or more sequences of the same region of comparable quality were available for the same species, we chose the longest sequence. Our final dataset consisted of data from 21 gene regions and a single accession from each of 623 species, following the nomenclature of Govaerts et al. (2007; SI Table 1). In our supermatrix, 98.6% of species are represented, at minimum, by at least one of the core gene regions or *ITS*, and 70% are represented by at least 2 of these regions, following the "scaffolding" recommendations of Hinchliff and Roalson (2013).

2.2. Phylogenetic analysis and molecular dating

We aligned sequences from each gene region using MUSCLE (Edgar, 2004) and subsequently edited the sequences manually in Geneious v.7.1.7 (Biomatters, available from http://www.geneious.com). We identified the optimum partition scheme using PartitionFinder v. 1.1.1 (Lanfear et al., 2012), and conducted Maximum Likelihood (ML) analysis on the partitioned dataset using default settings in RAxML v 8.0.9 (Stamatakis, 2014) as implemented in the CIPRES Science Gateway (Miller et al., 2010). This included 400 rapid bootstrap replicates and ten fast ML searches followed by ten thorough optimizations to identify the most likely tree and associated nodal support.

The best tree from the ML search was transformed to ultrametric using the penalized likelihood algorithm in treePL (Smith and O'Meara, 2012), which is better suited for large phylogenetic trees than BEAST (Drummond et al., 2012). We selected a smoothing parameter of 10 following the cross-validation approach and χ^2 test as implemented in treePL, in which seven values from 0.001 to 1000 were tested. We offset the minimum ages of nine nodes (Table 1) according to dates in a recent family-wide analysis (Spalink et al., submitted for publication). In that analysis, Cyperaceae was placed in a Poales-wide context, calibrated using 15 fossils and four secondary priors throughout Poales, and analyzed using BEAST (Drummond et al., 2012). The resulting BEAST chronogram was largely consistent with other dated phylogenies involving Cyperaceae (e.g., Escudero and Hipp, 2013; Viljoen et al., 2013), but provides additional dates for nodes of interest in this present study. Finally, we eliminated tips in the phylogeny for which distribution, climate, or soil were unavailable. Throughout this paper we will refer to this chronogram as the "species tree."

2.3. Distribution data

We mined the Global Biodiversity Information Facility (GBIF; www.gbif.org) for distribution data for Cyperaceae species in Mexico, Canada, and the United States, and downloaded all georeferenced samples recorded as having no known coordinate ambiguities. To supplement this dataset, we searched for online herbarium databases in North America that have not contributed to GBIF and obtained their specimen information. We also obtained specimen data by directly contacting curators at herbaria which have been databased but whose data are not available online. Using these approaches, we obtained additional data from the Consortium of Northeastern Herbaria (neherbaria.org), the Intermountain Regional Herbarium Network (intermountainbiota.org), KSC, SASK, USCH, and WIS. Once assembled, we culled the distribution database to remove duplicate records, accessions that were clearly outside of known species ranges, accessions with ambiguous taxonomy, and taxa that were not represented in the species tree.

Table 1
Age offsets for penalized likelihood dating. Minimum and maximum offsets are based on ages recovered in the worldwide Cyperaceae analysis of Spalink et al. (submitted for publication).

Node	Minimum offset (Ma)	Maximum offset (Ma)
Root	77	88
Cladium crown	26	36
Scleria crown	38	48
Eleocharis crown	31	41
Fimbristylis crown	25.7	35.7
Cyperus crown	23	33
Scirpus crown	28.7	38.7
Carex crown	20	30
Caricoid *Carex* crown	15	17
Core *Carex* crown	14	18

2.4. Transforming distribution into a continuous variable

To transform species distribution into a continuous variable, we first created a map of North America composed of 100 quadrats, each $6^\circ \times 6^\circ$, in QGIS 2.0.1-Dufour (Quantum GIS Development Team, 2014). We then overlaid the species' distributions onto this map to create a binary (presence and absence) matrix for each species in each quadrat using the R package maptools (Bivend and Lewin-Kof, 2014). We subsequently eliminated quadrats that contained no specimen records, which were almost exclusively in the northern portions of Nunavut and the Northwest Territories of Canada, and retained a total of 90 quadrats. Finally, we conducted correspondence analysis, an ordination technique similar to principal component analysis but better suited for discrete characters (Hirschfeld, 1935; Hill, 1974; Greenacre and Vrba, 1984; Greenacre, 2007), and retained the first five canonical axes for comparative phylogenetic analyses. Hereafter, we will refer to these traits as "continuous species distribution", or "CSD", in order to distinguish them from general discussions of species distribution in the traditional sense. Correspondence analysis was conducted using the dudi.coa function in the ade4 package (Dray and Dufour, 2007) in R (R Development Core Team, 2013).

2.5. Geography cladogram and community phylogenetics

To test the extent to which evolutionarily related species are geographically related, we utilized an approach conceptually similar to that commonly implemented in Parsimony Analysis of Endemicity (PAE; Morrone, 2014), which is used to identify geographic regions composed of congruently endemic species. In PAE, all molecular data are omitted, geographic areas are treated as tips, and the presence/absence of the species within the areas are treated as the characters used to build the cladogram. In this study, however, we are more interested in the relatedness of species based on their distributions than in the relatedness of areas based on their species compositions. For this reason, we instead treated the species as tips and geographic quadrats as characters; the presence or absence within each quadrat was treated as the character state for each species. In this context, species "relationships" are based on the similarity of their geographical distributions rather than shared evolutionary history. To construct a quadrat matrix for this analysis, we used the same approach as described in Section 2.4 but increased the resolution by using quadrats 1.68° x 1.68°. After overlaying species' distributions onto these quadrats and removing empty regions, 997 quadrats were retained. We then subjected this binary matrix of species' presence and absence to phylogenetic analysis to construct what we will refer to as the 'geography tree'. Comparing the species tree with the geography tree is a useful means to visualize the extent to which evolutionarily related species (i.e., species within a clade

in the species tree) are also geographically related (i.e., occurring within the same "clade" in the geography tree).

The geography tree was constructed using FastTree v2.1 (Price et al., 2010) as implemented in Geneious, following Lewis's (2003) recommendations for the use of a likelihood optimality criterion in the phylogenetic analysis of discrete characters. FastTree approximates a maximum likelihood tree, estimates branch lengths based on a simple Jukes–Cantor model, and calculates nodal support based on Shimodaira–Hasegawa tests of 1000 resampled datasets. In the context of this analysis, the tree topology reflects the geographical similarity of species, and branch lengths are proportional to the number of quadrats shared among taxa. We utilized an "empty" area, in which all taxa were coded as absent, as the outgroup. Using this geography tree, we identified "clades" composed of a minimum of five taxa and with nodal support of at least 70%, and characterized the distributions of the species within to define geographic regions united by floristic similarity (i.e., floristic regions). We will hereafter refer to these "clades" as "geoclades" to distinguish them from discussions of clades in the true phylogenetic sense.

We subsequently analyzed each of these floristic regions using our species tree under a community phylogenetic framework following Webb et al. (2002). We calculated the mean pairwise phylogenetic distance between all species residing in each floristic region, and compared the extent of phylogenetic clustering measured to that expected if the species were randomly distributed among the communities. We tested for significant departure from the null model of random distribution by simulating and measuring pairwise distance in 1000 permuted datasets, thereby determining whether species within a given geographic region are more closely related to each other than would be expected by chance. These analyses were conducted using the ses.mpd function in the R package picante (Kembel et al., 2010).

2.6. Characterizing climatic and edaphic niche

To characterize climatic and soil preferences for each species, we first downloaded raster layers at 30″ resolution for 19 bioclimatic variables from WorldClim (Hijmans et al., 2005). To characterize soil, we downloaded six continuous variables at the same spatial resolution from the Harmonized World Soils Database 1.21 (HWSD; FAO, 2012). These included: % topsoil clay, % topsoil gravel, % topsoil sand, % topsoil silt, % topsoil organic carbon, and topsoil pH. We then overlaid the species distribution data on each raster layer to obtain the values for each specimen accession using the extract function in the R package raster (Hijmans, 2014). We tested for autocorrelation among the 26 variables, and identified those that were over 70% correlated. We then eliminated the variables that were autocorrelated with at least two other variables. Using this approach, we were able to remove all instances of correlation greater than 70%. We retained Bio1, Bio 2, Bio 4, Bio 4, Bio 5, Bio 8, Bio 13, Bio 14, Bio 18, Bio 19, and all soil variables, and computed the median value and variance of each variable for each species, following Joly et al. (2014a).

2.7. Analysis of phylogenetic signal in range and niche traits

To determine if species' ranges and niches exhibit a phylogenetic signal comparable to what would be expected under a Brownian model of evolution, we calculated Blomberg's *K* (Blomberg et al., 2003) for the five CSD axes, median elevation, and the bioclimatic and edaphic variables. To further partition phylogenetic signal associated with species distribution, we also calculated *K* values for species' median latitudes and longitudes. We used 1000 datasets simulated under BM to test the hypothesis that the observed *K* value obtained for each trait is different than the expected value if the trait followed a Brownian pattern. We also tested the null hypothesis that these traits exhibit no more phylogenetic signal than would be expected if the values were randomly assigned to tips by comparing observed *K* values to the expected *K* values of 1,000 randomly permuted datasets. Using this approach, if traits exhibited less phylogenetic signal than would be expected under BM, we were able to detect if there was still more signal than if these traits were randomly distributed (Harmon-Threatt and Ackerly, 2013). All analyses were conducted using the R package phytools 0.4-31 (Revell, 2012).

2.8. Rates of speciation and evolution of range and niche

We measured rates of speciation as well as rates of evolution in range (CSD 1-5, mean latitude, mean longitude, and elevation), mean annual temperature (Bioclim 1), mean temperature seasonality, (Bioclim 4), mean precipitation during the warmest and coldest quarters (Bioclim 18 and 19, respectively), and soil (soil composition and pH variables) using BAMM v2.0 (Rabosky et al., 2014a, 2014b). For each analysis, we initiated two chains of 50,000,000 generations each, and assessed effective sampling and convergence using the R package CODA (Plummer et al., 2006), identifying the best-fitting models using Bayes factors.

In this study, we are exclusively interested in the dynamics of the North American sedge flora. Because the North American sedges are not monophyletic in the context of Cyperaceae worldwide, we do not make any claims about the net rates of speciation, range, and niche evolution of the family as a whole. Rather, we are interested in how these rates vary relative to each other within the context of the North American sedge flora. Therefore, to account for missing species in our phylogeny, we assigned all tips to genera and specified the proportion of each genus, from the North American flora only, that was sampled. In the genus *Carex*, we assigned tips to lower taxonomic levels to which we could confidently assign richness values. These included *Carex* subgenus *Vignea*, the Caricoid clade, and the Core *Carex* clade. Using this approach, all tips were assigned to a total of 21 clades. Clade assignments and associated species richness are reported in SI Table 2. Ultimately, we believe that the results from these analyses are conservative. Because these lineages have dispersed into and out of North America throughout their diversification, we would expect even greater rates of range and niche evolution than would be captured in these analyses.

2.9. Phylogenetic generalized least squares (PGLS) regressions

To determine which factors may be associated with rates of speciation, to explore correlations to species richness, and to identify geographic patterns of diversity, we conducted a series of phylogenetic generalized least squares (PGLS) regressions. These analyses were conducted on a phylogeny that was pruned to include one tip per clade, using the 21 clades identified in the BAMM analysis (SI Table 2) but with the inclusion of the monotypic genus *Amphiscirpus* Oteng-Yeb in the *Scirpus* L. clade. In a pairwise fashion, we regressed the following traits: species richness, range size (defined as the average number of quadrats inhabited by species in a clade), clade stem age, median and variance of latitude, longitude, elevation, mean annual temperature (Bioclim 1), temperature seasonality (Bioclim 4), precipitation during the warmest and coldest quarters (Bioclim 18 and 19, respectively), mean rates of speciation, rates of evolution of all geographic variables (including CSD), and rates of evolution of the soil and four bioclimatic variables listed above. Analyses were conducted twice, such that each trait was treated both as response and predictor for every pairwise comparison. All variables were log-transformed prior to regression. Variances were treated as the average of the variances found

among accessions of each species in the clade, as opposed to the total variance of all accessions of all species within the clade. All regressions were conducted using the pgls function in the R package caper (Orme et al., 2013).

3. Results

3.1. Supermatrix analysis

Our final supermatrix was composed of 623 species, 21 gene regions, a total concatenated alignment length of 20,595 base pairs, and ca. 80% missing data. Because nearly all taxa were represented by at least one of five primary gene regions, this level of missing data is substantially lower than many supermatrix analyses and should not lead to significant topological inaccuracies (Wiens and Morrill, 2011; Wiens and Tiu, 2012; Hinchliff and Roalson, 2013; Deng et al., 2015). The ML phylogeny (SI Fig. 1) exhibits a topology largely consistent with previous analyses, with most deviations involving clades that are poorly supported in all family-wide analyses to date (e.g., Muasya et al., 2009; Hinchliff and Roalson, 2013). One exception is in tribe Abildgaardieae, where we did not recover either *Fimbristylis* Vahl or *Bulbostylis* Kunth as monophyletic genera. Rather, *Bulbostylis* formed a grade sister to *Fimbristylis*.

The penalized likelihood chronogram (SI Fig. 2) places the crown age of the North American Cyperaceae at about 82.5 million years ago (Mya), the *Eleocharis* crown at 25 Mya, *Cyperus* crown at 22 Mya, and the *Carex* crown at 21 Mya, which is consistent with all previous studies to date (Escudero and Hipp, 2013; Viljoen et al., 2013).

3.2. Species distribution and correspondence analysis

We compiled a dataset of 310,166 georeferenced samples of North American sedges. We removed over 35% of the records because they were listed with outdated and ambiguous taxonomy, were clearly outside of known species ranges, or were duplicate records. Our culled distribution dataset consisted of 201,104 unique records (Fig. 1). Our final comparative dataset, which included species represented in the phylogeny and which also had distribution, climate, and soil data, was comprised of 547 species.

Mapping the georeferenced samples onto a map of the 997 quadrats, we found that the most diverse quadrat contained 243 species of Cyperaceae (Fig. 1A). This quadrat encompassed an area including the eastern portion of Long Island, NY and southern Connecticut. There appear to be six geographic regions within North America with exceptional species richness, loosely defined here as five or more contiguous quadrats each with 75 or more species. The first constitutes most of the northern Mid-Atlantic and New England, ranging from New Jersey in the South to Nova Scotia in the north, and extending west until eastern Ohio (Fig. 1A, outlined in red). A second area of richness is centered in the north Midwest, encompassing Wisconsin and Michigan in the north, and northern Illinois in the south (outlined in blue). This area is separated from the third area of species richness, which includes much of Missouri, northern Arkansas, and southern Illinois (outlined in green). Fourth, sedges appear concentrated along the north coast of the Gulf of Mexico, including much of Alabama (outlined in purple). A fifth area of species richness encompasses much of Colorado, northern New Mexico, eastern Utah, and southern Wyoming (outlined in yellow). Finally, a sixth area of concentrated sedge richness includes much of California, Oregon, Washington, western Idaho, and southern Alberta. The Cyperaceae generally appear to have less diversity in Nevada, Texas, Baja California, and much of Nunavut and northern Quebec. Although these patterns of richness tend to exhibit close proximity to major universities and herbaria, suggesting the potential of collection biases, these centers of sedge diversity are well documented in many regional and national floristic treatments (e.g., Gleason and Cronquist, 2003; Voss and Reznicek, 2012; Kartesz, 2013).

Quadrats that intersect the 44°N parallel exhibit the highest species richness with an average of 80.4 species per quadrat (Fig. 1B). This is followed by quadrats that intersect the 42°N and 46°N parallels, which have on average 79.9 and 74.5 species,

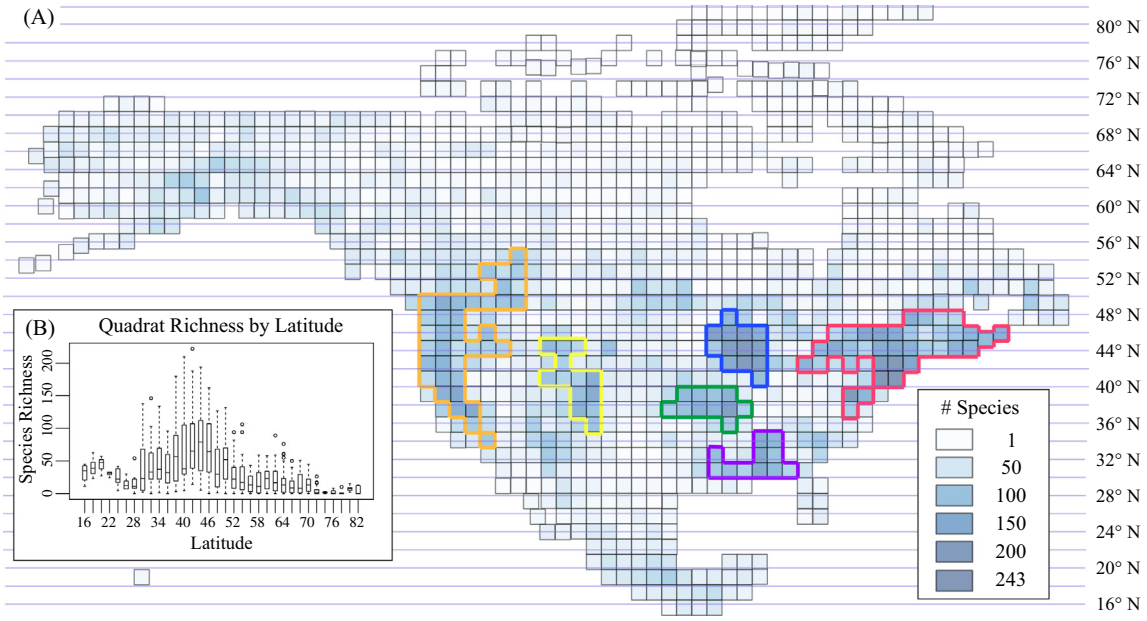

Fig. 1. Quadrat map of North America. The locations of georeferenced specimens were overlaid onto 997 quadrats in order to characterize species distributions of all sedges of North America, and to construct the geography tree in Fig. 2. (A) Species richness per quadrat. The shade of the quadrat represents the number of species within, as defined by the inset key. Regions with colored outlines consist of five or more contiguous quadrats containing at least 75 species. (B) Boxplot of latitudinal richness, defined as the mean number of species in each quadrat that are bisected by the lines of latitude in (A). 44°N bisects quadrats with the highest average species richness. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

respectively. Species richness drops substantially north of 50°N, with most quadrats having an average of 12–27 species between 50°N and 70°N, and fewer than 10 species between 72°N and 82°N.

The first five canonical axes accounted for nearly 50% of the variance associated with distribution (as compared to 8%, when the higher resolution matrix of 997 quadrats was utilized). Each axis captures different aspects of this variance, including both range size and region of distribution. Along the first axis species with low scores are endemic to the southeastern United States and those with high scores are widespread throughout the high latitudes, while along the second axis species with low scores are endemic to the northeast United States, those with medium scores are mostly western, and those with high scores are mostly southeastern. Along the third axis, species with low scores exhibit mostly small ranges and are eastern while those with high scores are mostly western endemics. The fourth axis ranges from southern endemics to western endemics, and the final axis ranges from Texas endemics to species largely restricted to the southeast United States.

3.3. Geography cladogram and community phylogenetics

Phylogenetic analysis of the quadrat matrix resulted in a geography tree with strong support for many species relationships based on their geographic similarity (Fig. 2A). Topological support along the backbone of this tree was weak, particularly among the three first diverging geographical lineages. We divided this cladogram into ten non-overlapping geoclades, each composed of a minimum of five species and with at least 70% nodal support, and characterized the distributions of the species within them. These geoclades differ not only in the geographic centers of distribution, but also in the number of quadrats that the species inhabit (Fig. 2C)

When treating the 10 geoclades as floristic regions in a community phylogenetic framework, six exhibited greater phylogenetic distance among co-occurring species than would be expected by chance (i.e., phylogenetic overdispersion) and four exhibited smaller phylogenetic distance among co-occurring species than would be exhibited by chance (i.e., phylogenetic clustering; Table 2). Regions exhibiting stronger phylogenetic clustering involved California, the Rocky Mountains, and high latitudes while regions with phylogenetic overdispersion generally involved the more eastern and southern areas (Table 2, Fig. 2). The two most overdispersed regions occur in the subtropical southeastern United States and Mexico (Fig. 2C vii) and in the arid southwest United States (Fig. 2C i).

3.4. Phylogenetic signal in range and niche variables

In all instances, observed K values were significantly lower than expected if these variables had evolved in a manner consistent with Brownian motion ($K < 1$; $p < 0.001$; Table 3). Topsoil % clay, % gravel, and % organic carbon exhibited no phylogenetic conservatism, while all other variables exhibited greater levels of conservatism than would be expected if they were randomly distributed throughout the phylogeny (Table 3). Species latitude was the most phylogenetically conserved trait, with K values over three times greater than species longitude, and the precipitation variables (Bioclim 12–19) were typically less conserved than the temperature variables (Bioclim 1–11).

3.5. Rates of speciation and evolution of range and niche

In our BAMM analyses, convergence of independent chains and effective sample sizes were achieved in excess of 200 after 50,000,000 generations for each variable. We uncovered five significant increases in speciation rate, all within the past 20 million years (Fig. 4A). These occurred within *Eleocharis, Cyperus,* the Core *Carex* clade, along the stem lineage of *Carex* subgenus *Vignea*, and the stem lineage of *Carex* sect. *Ovales* within subgenus *Vignea.* Comparable significant shifts in speciation rate involving the same clades have been identified in previous globally sampled analyses of Cyperaceae (Escudero et al., 2012; Escudero and Hipp, 2013). In general, our analyses recover these shifts as having occurred on more internal, recent nodes, as would be expected in our explicitly North American community based sampling where many earlier diverging lineages, found outside North America, are omitted. Ultimately, we suspect that a completely sampled, worldwide phylogeny of Cyperaceae would place significant shifts in diversification rate somewhere between those recovered in previous analyses and this current study. However, because of our sampling bias we limit discussion of these rates only as they pertain to rates of geographical and niche evolution.

For the species distribution and niche variables, we recovered multiple significant shifts in evolutionary rates (Fig. 4B–X). Most of these occurred within or including the same clades involved in shifts in speciation rate. We additionally uncovered significant increases in evolution rates of Bioclim 2 and 5 (mean diurnal temperature range and maximum temperature of warmest month, respectively) within the mostly boreal/arctic *Eriophorum*, of longitude at the most recent common ancestor of *Trichophorum*, *Scirpus*, and *Carex*, and of CSD Axes 1 and 4 in the Caricoid clade of *Carex*.

3.6. PGLS regressions among niche variables, range, and evolutionary rates

Phylogenetic generalized least squares regressions resulted in multiple significant correlations between the variables examined (Fig. 4). Of the 21 clades defined on the species tree (SI Table 2), those that are widespread are associated with lower temperatures, less precipitation in colder months (Bioclim 19), more temperature seasonality (Bioclim 4), and are more likely to occur in eastern North America. Clades that require more summer precipitation (Bioclim 18) are generally more species rich, more eastern in distribution, tend to have narrow ranges, and and occur at lower elevations. Clades that require more winter precipitation (Bioclim 19) are more western in distribution, occur over smaller latitudinal and longitudinal gradients, in areas with less temperature seasonality, and require warmer annual temperatures. Clades that occur in higher elevations also exhibit more variance in elevation and less summer precipitation. The only bioclimatic variable associated with clade richness is summer precipitation, which was positively correlated. However, clade richness was very strongly correlated to rates of speciation and rates of range (including CSD, latitude, longitude, and elevation), climate, and soil evolution. The evolutionary rates of most range parameters were also correlated with rates of niche evolution, and both range and niche parameters were strongly correlated to rate of speciation.

4. Discussion

The evolution of geographical "place" and climatic and edaphic niche "space" is fundamental to lineage diversification (Vamosi and Vamosi, 2011; Loera et al., 2012), but is difficult to both quantify and analyze in a phylogenetically rigorous fashion. In this paper we have proposed two methods for characterizing geography that are conducive to phylogenetic analysis in a continuous framework. We use these methods to test key hypotheses regarding the evolution of North American Cyperaceae, namely that the diversification of the North American sedges is associated with both range and niche evolution and that geographical diversification involves adaptation to new ecological conditions.

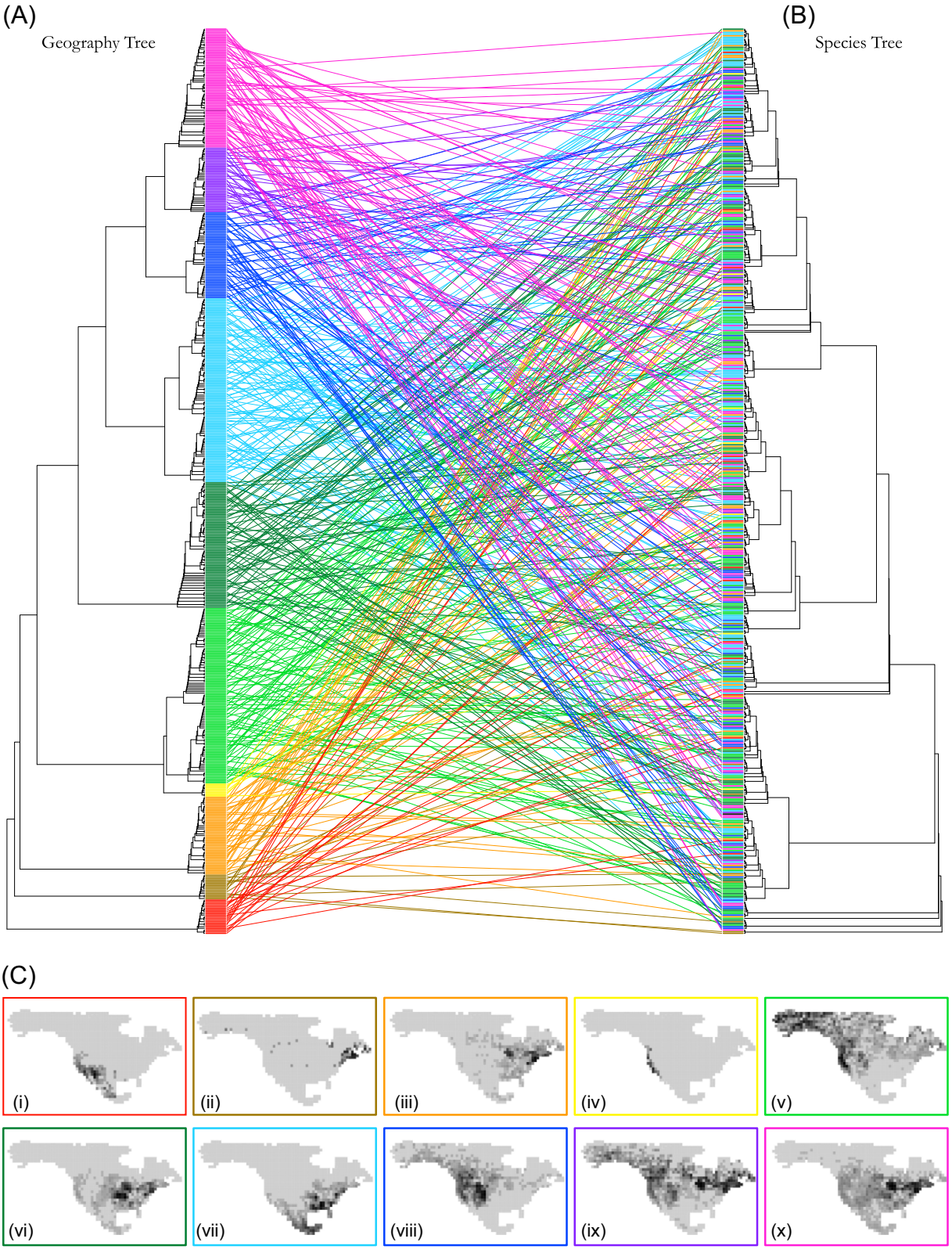

Fig. 2. Topological comparison of geography tree and species tree of North American sedges. (A) Geography tree, constructed using a phylogenetic analysis of distribution data of all North American sedges. Tip labels correspond to geoclades consisting of at least five species and support values of at least 70%. The distributions of species within these geoclades are characterized by the maps in (C), with the tip label color corresponding to the map outline color. (B) Species tree constructed using a phylogenetic analysis of supermatrix molecular data for the same species as in (A). Tip labels on the species tree match the colors of the label in the geography tree corresponding to the same species, the location of which is indicated by the lines connecting the two trees. (C). Distribution of species within each of the geoclades in the geography tree. Within each map, species richness is indicated by the shade of the highlighted quadrats, with darker colors indicating higher relative diversity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Mean phylogenetic pairwise distance among floristic regions. The floristic regions correspond to those identified in Fig. 2. The Richness column depicts the number of species occurring in each of the floristic regions. Observed mean pairwise distances (MPD), and those calculated from 1000 permuted datasets, are presented. Significant *p*-values ($p < 0.05$) suggest that the species within the floristic region are more closely related to each other than would be expected by chance.

	Richness	Observed MPD	Simulated MPD	*P* value
Region 1	21	94.46	75.40	0.95
Region 2	14	69.77	74.20	0.36
Region 3	47	77.38	74.67	0.65
Region 4	8	30.80	74.64	0.01
Region 5	106	50.08	74.83	0.00
Region 6	76	74.63	74.74	0.49
Region 7	111	97.96	74.43	1.00
Region 8	52	31.51	74.17	0.00
Region 9	40	60.74	74.70	0.03
Region 10	72	78.10	74.74	0.74

Table 3
Measures of phylogenetic signal and tests of significance. Observed values of Blomberg's *K* for the continuous species distribution (CSD) axes, mean latitude, longitude, and the climatic and edaphic variables. pr{Brownian} and pr{Random} columns depict the probability that the observed *K* values fit expected values under a Brownian model of evolution or are randomly distributed. In all instances, significant *p* values ($p < 0.05$) indicate "less phylogenetic signal than Brownian Motion" and "more phylogenetic signal than random." Abbreviations: Bioclim 1 = mean average temperature; Bioclim 2 = mean diurnal range; Bioclim 4 = temperature seasonality; Bioclim 5 = max temperature of warmest quarter; Bioclim 8 = mean temperature of wettest quarter; Bioclim 13 = precipitation of wettest month; Bioclim 14 = precipitation of driest month; Bioclim 18 = precipitation of warmest quarter; Bioclim 19 = precipitation of coldest quarter.

Variable	*K*	pr{Random}	pr{Brownian}
CSD Axis 1	0.068	0.001	0.001
CSD Axis 2	0.079	0.001	0.001
CSD Axis 3	0.060	0.001	0.001
CSD Axis 4	0.029	0.002	0.001
CSD Axis 5	0.030	0.022	0.001
Latitude	0.101	0.001	0.001
Longitude	0.031	0.002	0.001
Elevation	0.039	0.001	0.001
Bioclim 1	0.093	0.001	0.001
Bioclim 2	0.034	0.002	0.001
Bioclim 4	0.051	0.001	0.001
Bioclim 5	0.088	0.001	0.001
Bioclim 8	0.045	0.001	0.001
Bioclim 13	0.033	0.001	0.001
Bioclim 14	0.023	0.013	0.001
Bioclim 18	0.041	0.001	0.001
Bioclim 19	0.023	0.024	0.001
% Clay	0.020	0.077	0.001
% Gravel	0.015	0.509	0.001
% Sand	0.031	0.003	0.001
% Silt	0.022	0.016	0.001
% Organic carbon	0.014	0.75	0.001
Soil pH	0.031	0.001	0.001

We incorporate data from over 200,000 specimens and over 20,000 base pairs of molecular sequence data, which to our knowledge is the largest study of its kind. Our analyses consistently support the hypotheses that the diversification of the North American sedge flora is associated, at least in part, with the evolution of species distributions and climatic and edaphic niches, and that the exploration of geographic "place" is correlated to that of niche "space". Comparisons of the species tree and geography tree demonstrate substantial incongruence between the two, which are confirmed with community phylogenetic metrics. Tests of phylogenetic signal indicate that geography and niche are more divergent than if they had evolved in a Brownian fashion, that shifts in distribution occur more commonly along a longitudinal than latitudinal gradient, and that temperature tolerance is more limiting than precipitation or soil tolerance. Lastly, measurements of evolutionary rates imply that geographical diversification and speciation are very strongly correlated and that the evolution of geographical place corresponds strongly with exploration of niche space.

4.1. Evolution of range and niche is associated with species diversification

On a global scale, geography has played a critical role in the diversification of Cyperaceae, with both long-distance dispersal and sympatric speciation commonly invoked when geography is coded as broad, discrete regions (Escudero et al., 2009; Viljoen et al., 2013; Villaverde et al., 2015). Our analyses of the North American sedge flora, however, reveal very few instances of sympatric speciation as would be recovered through ancestral range estimations of discretized space in programs such as Lagrange (Ree and Smith, 2008) or BioGeoBEARS (Matzke, 2013, 2014). Instead, treating distribution in a continuous manner reveals that a majority of the sympatric speciation events uncovered in such previous analyses may in fact be allopatric, resulting either from dispersal or vicariance within a single land mass.

This is perhaps most apparent visually when comparing the geographic (Fig. 2A) and phylogenetic relationships of the sedges (Fig. 2B), as species composing the various clades in the species tree appear to be widely distributed among geoclades in the geography tree. When treating the geoclades as floristic regions in a community ecology sense and explicitly measuring their phylogenetic structure, we find that 60% of these regions are significantly phylogenetically overdispersed (Fig. 2C, Table 2; Webb et al., 2002). Thus, closely related species tend to be geographically unrelated, and distantly related species appear to have converged geographically onto these areas. This pattern is not found, however, in either mountainous or high latitude regions of North America (Fig. 2c vi, v, xii, ix), which instead exhibit significant phylogenetic clustering. Transitions into cold regions, such as those at high latitudes and elevations, are relatively rare among the "major ecological transitions" (Edwards and Donoghue, 2013), and thus we might expect greater levels of niche conservatism and phylogenetic clustering in these areas. Our measurements of phylogenetic signal confirm these patterns of community of assemblage. Indeed, species' latitude was the single most conserved variable that we analyzed (Table 3), with a *K* value over three times greater than species' longitude. Similarly, the temperature variables exhibit more conservatism than either precipitation or soil (Table 3). Thus, lineage splitting is more commonly associated with longitudinal (east–west) than latitudinal (north–south) movement, and far more commonly associated with shifts in precipitation and soil regimes than in temperature regimes.

Measuring rates of speciation and range and niche evolution further clarify these patterns. In all cases, significant shifts in speciation rate occur at the same time as, or soon after, significant shifts in rate of range, elevation, temperature, precipitation, or soil evolution, providing strong evidence in support of the hypothesis that the diversification of the North American sedge flora is associated with ecological and geographical divergence (Ackerly, 2009). In some instances, these shifts in rate of evolution occur at the same moment, as with the accelerations in speciation and CSD Axis 4 in *Carex* sect. *Ovales,* or in speciation and winter precipitation in *Carex* subgenus *Vignea.* In other instances, significant shifts occur at different times within the same clade. For example, within *Eleocharis,* increases in rates of evolution of elevation, longitude, CSD Axis 4, and summer temperature occur prior to shifts in speciation rates. Furthermore, in nearly all cases significant shifts in rate of species distribution evolution are associated with significant shifts in rate of niche evolution, suggesting that the exploration of geographic place corresponds very closely to exploration of niche space.

Contrary to theoretical expectations (Givnish, 2010; Vamosi and Vamosi, 2011), our analyses found no correlation between range size and richness (Fig. 4). In a study of angiosperm families, Vamosi and Vamosi (2011) found that the area available to a lineage (characterized as the sum of the area of ecoregions, on a global scale, in which the family resides) is the single greatest predictor of that families' species richness. The sedges, however, do not fit that trend, as widespread clades can be either hyperdiverse (e.g., *Carex*) or relatively species poor (e.g., *Scirpus*). We also did not find any association between clade age and cold tolerance or between most of the climate traits and clade richness, which are both patterns identified in North American tree lineages (Hawkins et al., 2014; Qian et al., 2014). Rather, with the exception of a weak positive correlation to the amount of summer precipitation, the only traits related to clade richness in the North American Cyperaceae are rates of speciation and rates of geographical and ecological diversification. That is to say, species rich clades are diverse primarily because they underwent rapid speciation and rapid geographical and ecological evolution. Finally, geographic and ecological rates of evolution are strongly and positively correlated with each other. Rates of geographic and edaphic evolution are more strongly correlated with rates of speciation than any other trait. These results lend strong support to the hypotheses that species rich sedge clades have exhibited rapid rates of speciation, have experienced rapid geographical diversification, and have rapidly explored new niche space in the process.

4.2. Latitudinal patterns of diversity within Cyperaceae

The Cyperaceae is one of the few angiosperm families that exhibits higher species richness in temperate vs. tropical regions (Escudero et al., 2012; Kerkhoff et al., 2014; Tang et al., 2014). Escudero et al. (2012) suggest that the inverse latitudinal richness gradient in *Carex* may have been driven in part by significant cooling beginning in the Oligocene (~33.9 Ma) and again in the Pliocene (~2.6 Ma). During these times, higher latitudes would have experienced pronounced cooling, increasing the presence of boreal niche space available for colonization in these lineages (Graham, 1999, 2011, 2012; Zachos et al., 2001.). We further suggest that these patterns of latitudinal species richness may be the result of events during the Last Glacial Maximum (LGM), where areas of highest diversity correspond to the southernmost extent of the ice sheets (Fig. 1; Jackson et al., 2000; Clark and Mix, 2002; Marshall et al., 2002; Clark et al., 2009). These areas are now rich in both open wetlands and forested uplands, habitats where sedges exhibit peak diversity. On the other hand, areas north of 50°N were entirely glaciated during the LGM and colonization of these areas is necessarily a much more recent phenomenon.

Adaptation to boreal and arctic climes appears to be uncommon among sedge lineages, and lineages that do adapt to these biomes tend to exhibit niche conservatism and lower rates of diversification (Table 3, Fig. 4; Escudero et al., 2012; Gebauer et al., 2014). These patterns are reflected in other boreal or arctic lineages as well. For example, Tkach et al. (2014) found that arctic lineages of *Pedicularis* L. (Orobanchaceae) mostly evolved from high elevation ancestors, and that these exhibit high levels of niche conservatism. Similarly, Hawkins et al. (2014) found that adaptation to cold tolerance is a trait significantly associated with phylogenetic niche conservatism among North American forest trees. Generally, transitions to higher latitudes and colder climates are difficult to achieve (Smith and Donoghue, 2010; Edwards and Donoghue, 2013). Our results are consistent with these patterns, which suggest that the traits most closely associated with cold tolerance (i.e., latitude and temperature) are more conserved than precipitation, soil, and longitude (Table 3), that the most phylogenetically clustered sedge communities occur at high latitudes and elevations (Table 2, Fig. 2C), and that the diversification of high latitude sedge lineages (e.g., *Eriophorum*, *Carex* sections *Phacocystis* and *Vesicaria*) is closely associated with significant shifts in rate of temperature, precipitation, and soil evolution.

Ultimately, we believe that these patterns merit additional investigation, particularly regarding the timing of adaptation to higher elevations and latitudes, the lineages involved, and potential key innovations or functional traits that may have facilitated these transitions (Escudero et al., 2012; Gebauer et al., 2014). Ancestral state reconstructions of climate and soil preferences as well as ancestral range estimations would likely prove to be a very useful contribution to this field of inquiry, but would either require more thorough taxonomic sampling on a global scale or a narrowed focus on particular clades within the North American flora.

4.3. CSD or latitude and longitude?

In this paper we suggest the use of correspondence analysis to reduce a large binary matrix of species presences and absences to construct a continuous trait for comparative phylogenetic analysis. We complement this approach by subjecting the centroid latitudes and longitudes to the same suite of analyses. Both approaches have obvious advantages and disadvantages. For example, in the CSD approach, only a portion of the variance associated with species distribution is captured in the canonical axes. The same is arguably true in the centroid approach, however, for which measures of range shape and extent are entirely unaccounted. One clear advantage of the centroid approach is that latitudes and longitudes are inherently easier to visualize and discuss than axes constructed from dimension reduction approaches such as correspondence or principal component analyses.

We maintain that both approaches are useful, though, particularly when integrated with each other. Indeed, the information used in the centroid approach was captured in the CSD approach, as the rate of latitudinal evolution is correlated with rate of evolution of CSD Axes 1, 4, and 5 and the rate of longitudinal evolution is correlated with rate of evolution of CSD Axes 1 and 3 (Fig. 4). Furthermore, all of the traits associated with rates of latitudinal or longitudinal evolution are similarly associated with those of the CSD axes. The CSD axes, however, recovered information not captured by the centroids alone. For example, CSD Axis 2 was not correlated to either latitude or longitude. In addition, rates of CSD evolution were correlated with rates of evolution of Bioclim 1, 4, and 19, which the centroids were not. Finally, while a significant shift in diversification rate was recovered in the genus *Cyperus*, the only geographical traits with corresponding shifts in evolutionary rate were CSD Axes 4 and 5 (Fig. 3). We therefore suggest that the CSD approach captures more information than centroids alone; however, interpretation of these data are more challenging.

4.4. Conclusions and final remarks

Despite the importance of geography and niche to the diversification of the North American Cyperaceae as demonstrated by our analyses, these results would be strengthened by the inclusion of additional data. First, species sampling should be broadened to incorporate a truly monophyletic group of organisms. While global sampling of all 5000+ species of Cyperaceae, including both molecular data as well as complete distribution data, is unlikely to be accomplished in the near future, this would obviously be a superior dataset. In the meantime, additional sampling outside of North America would be useful to determine the role of transcontinental dispersal in diversification and to determine if patterns revealed in these analyses are reflected elsewhere. The work currently being conducted by the Global *Carex* Group (Pedro et al., 2014) to form a worldwide phylogeny of *Carex* will be a major step forward in

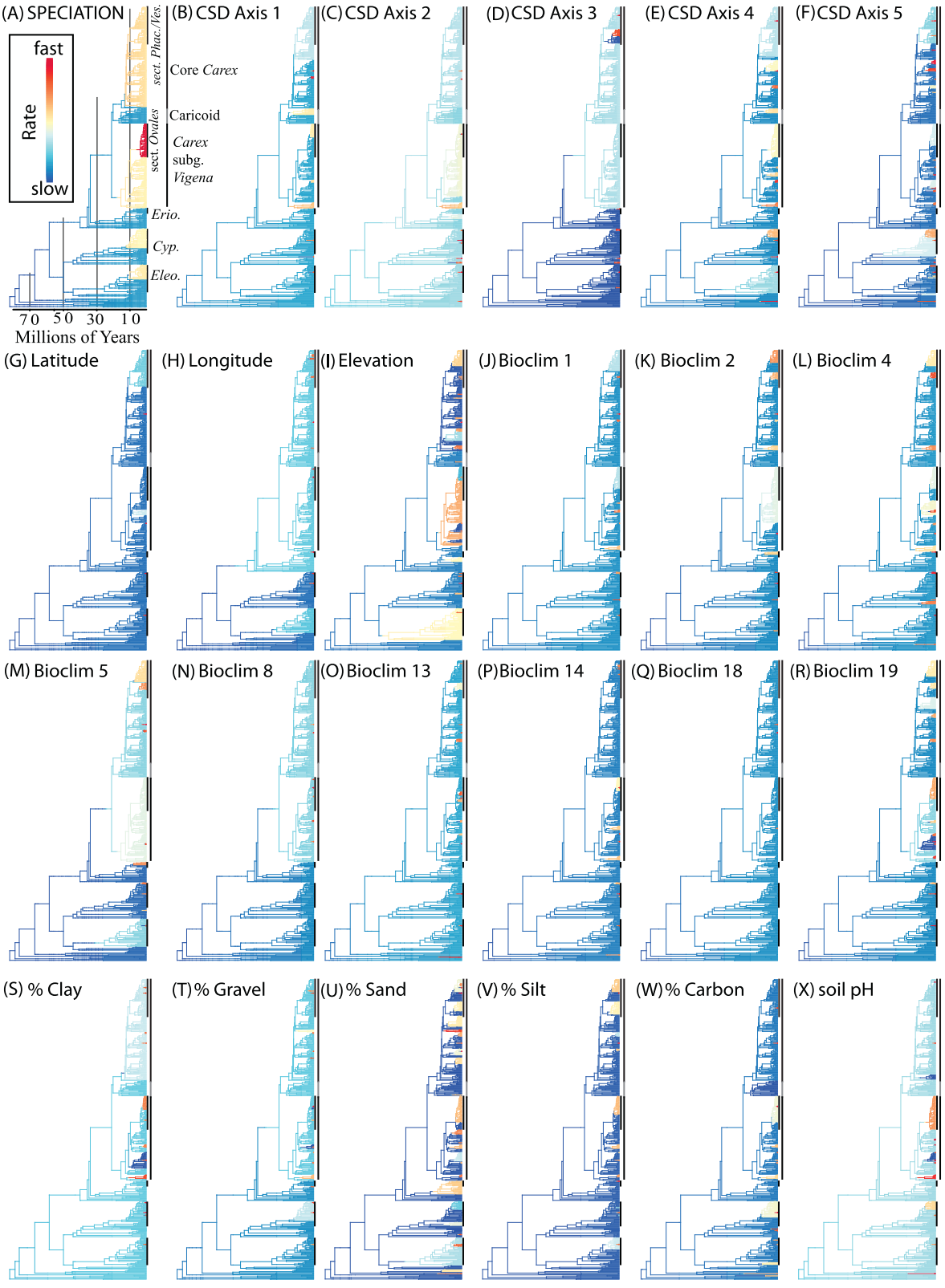

Fig. 3. Rates of evolution of North American sedges. In A–X, warm colors correspond to fast rates of evolution and cool colors to slow rates of evolution. All changes in branch color are significant shifts in rate of evolution. (A) Net rate of speciation. Chronogram of North American Cyperaceae, with branching times indicated by time-scale and branch colors reflecting net rates of speciation. Major clades specifically referred to in the text are labeled. B–X. Rates of evolution of continuous species distribution (CSD) axes, mean latitude, mean longitude, elevation, and temperature, precipitation, and soil parameters. Black/gray bars along the tips correspond to the labeled clades in Fig. 4A. Abbreviations: Eleo = *Eleocharis*; Cyp = *Cyperus*; Erio = *Eriophorum*; Phaco = *Carex* section *Phacocystis*; Ves = *Carex* section *Vesicaria*; CSD = continuous species distribution; Bioclim 1 = mean average temperature; Bioclim 2 = mean diurnal range; Bioclim 4 = temperature seasonality; Bioclim 5 = max temperature of warmest quarter; Bioclim 8 = mean temperature of wettest quarter; Bioclim 13 = precipitation of wettest month; Bioclim 14 = precipitation of driest month; Bioclim 18 = precipitation of warmest quarter; Bioclim 19 = precipitation of coldest quarter. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

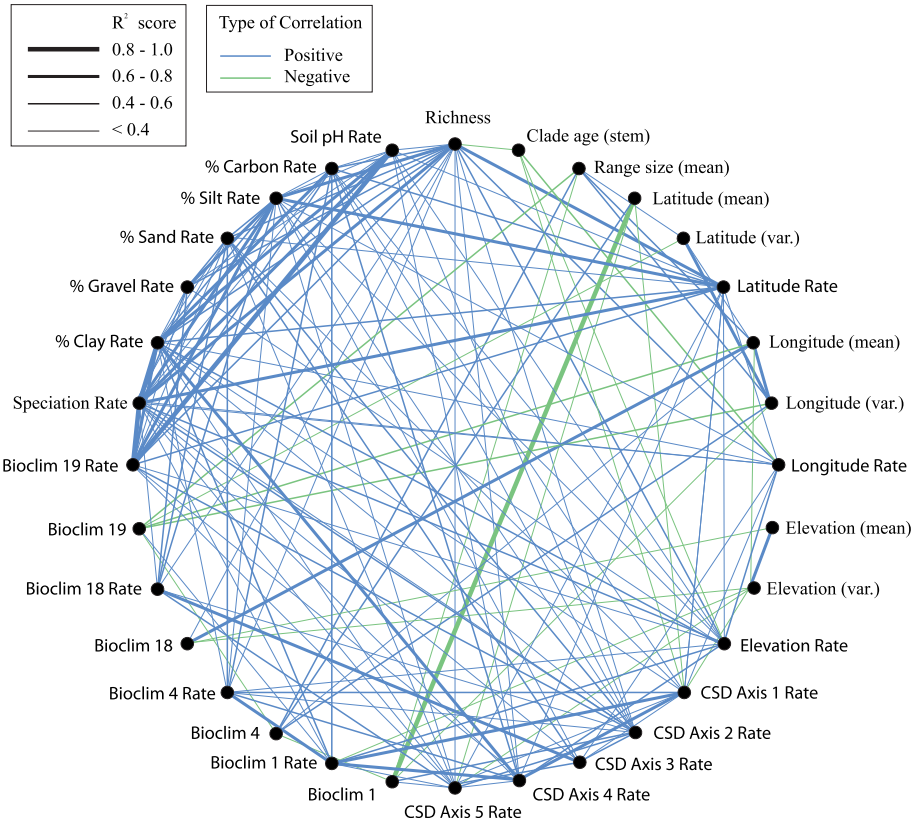

Fig. 4. Phylogenetic generalized least squares correlations among traits in the North American sedge flora. All variables listed were included in pairwise regressions. Statistically significant relationships between variables are depicted with colored connecting lines, with blue lines denoting positive correlations and green line denoting negative correlations. The thicknesses of the lines indicate the strength of the correlations, according to the inset key of R^2 scores. All rate variables were treated as average clade rates from BAMM analyses. Abbreviations: CSD = continuous species distribution; Bioclim 1 = mean average temperature; Bioclim 4 = temperature seasonality; Bioclim 18 = precipitation of warmest quarter; Bioclim 19 = precipitation of coldest quarter. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

this regard. Ecological niche modeling (e.g., Maxent; Philips and Dudík, 2008) could also prove a useful tool, particularly to predict distributions that may not currently be represented by available specimen data and to measure the extent of range overlap among species. We also suggest the inclusion of variables that may be contributing to reproductive isolation (e.g., chromosomal structure in the highly heterogeneous, agmatoploid *Carex*; Escudero et al., 2012), characters that may be related to dispersibility (e.g., fruit and seed morphology), that would facilitate exploration of niche space (e.g., photosynthetic pathway, adaptive traits facilitating latitudinal transitions), or are otherwise important functional traits. Projections of all of these traits into the past (i.e., through ancestral state reconstructions) would allow a more thorough exploration of how paleoclimates may have facilitated adaptation to temperature and precipitation regimes, particularly in relation to patterns of speciation (Töpel et al., 2012). This historical context may also be foundational to present conservation efforts, for example in aiding our ability to assess risk, predict future distributions, and identify suitable habitats as the global climate continues to change (Johnson et al., 2014; Lawing and Matzke, 2014). Finally, ongoing regional floristic barcoding projects, such as those currently ongoing in Wisconsin (Cameron et al., 2014), will continue to increase our ability to construct densely populated phylogenies in order to investigate geographical and ecological patterns of diversification across a broad spectrum of lineages, to identify common trends among these lineages, and to assess risks to long-term survival in a rigorous fashion (Davis et al., 2014; Joly et al., 2014b).

These analyses rely heavily upon data that are publicly and freely available in online repositories. We advise caution, however, in utilizing these data without first assessing their quality. In this paper, we omitted more than a third of all georeferenced samples from GBIF due to a variety of issues. Furthermore, data in online repositories are often incomplete (e.g., Drew, 2013; Drew et al., 2013; Hinchliff and Smith, 2014). While many herbaria and museums have contributed to repositories such as GBIF, the databasing of specimens in many collections is still in progress. This is particularly true for smaller, more local, regional collections, including teaching collections. We believe that these additional collections will be an invaluable resource for the development of more fine-scale, high resolution distribution datasets, and we encourage efforts to make these records available (Funk and Richardson, 2002; Pyke and Ehrlich, 2010; Lavoie, 2013; Rocha et al., 2014).

Acknowledgements

This work was funded by an NSF Dimensions of Biodiversity Grant to KMC, TJG, and KJS (DEB-1046355), and an NSF DDIG to KJS and DS (DDIG 1311153). Additional support was provided by an NSF graduate research fellowship, BSA and ASPT graduate student research grants, and a UW Department of Botany *Flora Aeterna* Fellowship and Davis Research grant to DS. We thank Donald Waller and two anonymous reviewers for helpful comments on the manuscript, and Amelia Krug for contributing to lab work. We also thank WIS for allowing destructive sampling of specimens for DNA extraction, and the Consortium of Northeastern Herbaria (neherbaria.org), the Intermountain Regional Herbarium Network (intermountainbiota.org), KSC, SASK, and USCH for providing specimen locality data. This work constitutes a portion of the research

conducted by DS to fulfill the requirements for a Ph.D. at the University of Wisconsin-Madison.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.09.028.

References

Ackerly, D., 2009. Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. Proc. Natl. Acad. Sci. USA 106, 19699–19706.

Andreasen, K., Baldwin, B.G., Bremer, B., 1999. Phylogenetic utility of the nuclear rDNA *ITS* region in the subfamily Ixoroideae (Rubiaceae): comparisons with cpDNA *rbcL* sequence data. Plant Syst. Evol. 217, 119–135.

Araya, Y.N., Silvertown, J., Gowing, D.J., McConway, K.J., Linder, H.P., Midgley, G., 2012. Do niche-structured plant communities exhibit phylogenetic conservatism? A test case in an endemic clade. J. Ecol. 100, 1434–1439.

Ball, P.W., Reznicek, A.A., Murray, D.F., 2003. Cyperaceae. In: Flora of North America Editorial Committee (Eds.), 1993+ Flora of North America North of Mexico. 18+ Vols. New York and Oxford., Vol. 23.

Bivend, R., Lewin-Kof, N., 2014. Maptools: Tools for Reading and Handling Spatial Objects. R package Version 0.8-30. <http://CRAN.R-project.org/package=maptools>.

Blomberg, S.P., Garland, T., Ives, A.R., 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57, 717–745.

Boucher, F.C., Thuiller, W., Davies, J.T., Lavergne, S., 2014. Neutral biogeography and the evolution of climatic niches. Am. Nat. 183, 573–584.

Cameron, K.C., Sytsma, K.J., Waller, D.H., Givnish, T.G., Pace, M.C., Li, P., Spalink, D., Drummond, C., Zaborsky, J., Rose, J., Alverson, B., Kriebel, R., 2014. A molecular phylogeny of the Wisconsin Flora. Abstract # 570, presented at Botany 2014, Boise, ID.

Clark, P.U., Mix, A.C., 2002. Ice sheets and sea level of the Last Glacial Maximum. Quat. Sci. Rev. 21, 1–7.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J. X., Hostetler, S.W., McCabe, A.M., 2009. The last glacial maximum. Science 325, 710–714.

Cornwell, W.K., Schwilk, D.W., Ackerly, D.D., 2006. A trait-based test for habitat filtering: convex hull volume. Ecology 87, 1465–1471.

Crisp, M.D., Cook, L.G., 2012. Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? New Phytol. 196, 681–694.

Davis, C.C., Schaefer, H., Ruhfel, B.R., Donoghue, M.J., Edwards, E.J., 2014. Climates and clades: biased methods, biased results. Available from: <arXiv:1406.5211 [q-bio.PE]>.

Deng, J., Drew, B.T., Mavrodiev, E.V., Gitzendanner, M.A., Soltis, P.S., Soltis, D.E., 2015. Phylogeny, divergence times, and historical biogeography of the angiosperm family Saxifragaceae. Mol. Phylogenet. Evol. 83, 86–98.

Donoghue, M.J., Edwards, E.J., 2014. Biome shifts and niche evolution in plants. Annu. Rev. Ecol. Evol. Syst. 45, 547–572.

Dray, S., Dufour, A.B., 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. 22, 1–20.

Drew, B.T., 2013. Data deposition: missing data means holes in tree of life. Nature 493, 305.

Drew, B.T., Sytsma, K.J., 2012. Phylogenetics, biogeography, and staminal evolution in the tribe Mentheae (Lamiaceae). Am. J. Bot. 99, 933–953.

Drew, B.T., Gazis, R., Cabezas, P., Swithers, K.S., Deng, J., Rodriguez, R., Katz, L.A., Crandall, K.A., Hibbett, D.S., Soltis, D.E., 2013. Lost branches on the tree of life. PLoS Biol. 11, e1001636. http://dx.doi.org/10.1371/journal.pbio.1001636.s006.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

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

Edwards, E.J., Donoghue, M.J., 2013. Is it easy to move and easy to evolve? Evolutionary accessibility and adaptation. J. Exp. Bot. 64, 4047–4052.

Escudero, M., Hipp, A., 2013. Shifts in diversification rates and clade ages explain species richness in higher-level sedge taxa (Cyperaceae). Am. J. Bot. 100, 2403–2411.

Escudero, M., Valcárcel, V., Vargas, P., Luceño, M., 2009. Significance of ecological vicariance and long-distance dispersal in the diversification of *Carex* sect. *Spirostachyae* (Cyperaceae). Am. J. Bot. 96, 2100–2114.

Escudero, M., Hipp, A.L., Waterway, M.J., Valente, L.M., 2012. Diversification rates and chromosome evolution in the most diverse angiosperm genus of the temperate zone (*Carex*, Cyperaceae). Mol. Phylogenet. Evol. 63, 650–655.

FAO, 2012. Harmonized world soil database (version 1.21). FAO, Rome and IIASA, Laxenburg.

Funk, V.A., Richardson, K.S., 2002. Systematic data in biodiversity studies: use it or lose it. Syst. Biol. 2, 303–316.

Gebauer, S., Starr, J.R., Hoffmann, M.H., 2014. Parallel and convergent diversification in two northern hemispheric species-rich *Carex* lineages (Cyperaceae). Org. Divers. Evol. 14, 247–258.

Geer, L.Y., Marchler-Bauer, A., Geer, R.C., Han, L., He, J., He, S., Liu, C., Shi, W., Bryant, S.H., 2010. The NCBI BioSystems database. Nucleic Acids Res. 38. http://dx.doi.org/10.1093/nar/gkp858.

Gilmour, C.N., Starr, J.R., Naczi, R.F.C., 2013. *Calliscirpus*, a new genus for two narrow endemics of the California Floristic Province, *C. criniger* and *C. brachythrix sp. Nov.* (Cyperaceae). Kew Bull. 68, 85–105.

Givnish, T.J., 2010. Ecology of plant speciation. Taxon 59, 1326–1366.

Givnish, T.J., Barfuss, M.H., Van Ee, B., Riina, R., Schulte, K., Horres, R., Gonsiska, P.A., Jabaily, R.S., Crayn, D.M., Smith, J.A.C., Winter, K., Brown, G.K., Evans, T.M., Holst, B.K., Luther, H., Till, W., Zizka, G., Berry, P.E., Sytsma, K.J., 2014. Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. Mol. Phylogenet. Evol. 71, 55–78.

Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J.D., Clements, M.A., Arroyo, M.T.K., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H., Cameron, K.M., 2015. Orchid phylogenomics and multiple drivers of their extraordinary diversification. Proc. R. Soc. B 282, 20151553.

Gleason, H.A., Cronquist, A., 2003. Manual of Vascular Plants of Northeastern United States and Adjacent Canada, Ed. 2. New York Botanical Garden.

Govaerts, R., Simpson, D.A., Goetghebeur, P., Wilson, K., Egorova, T., Bruhl, J.J., 2007. World Checklist of Cyperaceae. The Board of Trustees of the Royal Botanic Gardens, Kew.

Graham, A., 1999. Late Cretaceous and Cenozoic history of North American vegetation. Oxford University Press, New York.

Graham, A., 2011. The age and diversification of terrestrial New World ecosystems through Cretaceous and Cenozoic time. Am. J. Bot. 98, 336–351.

Graham, A., 2012. Sequencing New World ecosystems: comparisons of the Cretaceous and Cenozoic appearance of habitats with biome-characterizing plant groups. Ann. Mo. Bot. Gard. 98, 524–538.

Grandcolas, P., Nattier, R., Legendre, F., Pellens, R., 2011. Mapping extrinsic traits such as extinction risks or modeled bioclimatic niches on phylogenies: does it make sense at all? Cladistics 27, 181–185.

Greenacre, M.J., Vrba, E.S., 1984. Graphical display and interpretation of antelope census data in African wildlife areas, using correspondence analysis. Ecology 65, 984–997.

Greenacre, M.J., 2007. Correspondence Analysis in Practice, second ed. Chapman & Hall/CRC, Boca Raton.

Harmon-Threatt, A.N., Ackerly, D.D., 2013. Filtering across spatial scales: phylogeny, biogeography and community structure in bumble bees. PLoS ONE 8, e60446.

Hawkins, B.A., Rueda, M., Rangel, T.F., Field, R., Diniz-Filho, J.A.F., 2014. Community phylogenetics at the biogeographical scale: cold tolerance, niche conservatism and the structure of North American forests. J. Biogeogr. 41, 23–38.

Hijmans, R.J., 2014. raster: Geographic data Analysis and Modeling. R package version 2.3-12. <http://CRAN.R-project.org/package=raster>.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.

Hill, M.O., 1974. Correspondence analysis: a neglected multivariate method. J. Roy. Stat. Soc. C 23, 340–354.

Hinchliff, C.E., Roalson, E.H., 2013. Using supermatrices for phylogenetic inquiry: an example using the sedges. Syst. Biol. 62, 205–219.

Hinchliff, C.E., Smith, S.A., 2014. Some limitations of public sequence data for phylogenetic inference (in plants). PLoS ONE 9, e98986.

Hirschfeld, H.O., 1935. A connection between correlation and contingency. Math. Proc. Cambridge Philos. Soc. 31, 520–524.

Hua, X., Wiens, J.J., 2013. How does climate influence speciation? Am. Nat. 182, 1–12.

Jabaily, R.S., Sytsma, K.J., 2010. Phylogenetics of *Puya* (Bromeliaceae): placement, major lineages, and evolution of Chilean species. Am. J. Bot. 97, 337–356.

Jackson, S.T., Webb, R.S., Anderson, K.H., Overpeck, J.T., Webb III, T., Williams, J.W., Hansen, B.C.S., 2000. Vegetation and environment in eastern North America during the Last Glacial Maximum. Quat. Sci. Rev. 19, 489–508.

Johnson, A.L., Govindarajulu, R., Ashman, T., 2014. Bioclimatic evaluation of geographical range in *Fragaria* (Rosaceae): consequences of variation in breeding system, ploidy and species age. Bot. J. Linn. Soc. 176, 99–114.

Joly, S., Heenan, P.B., Lockhart, P.J., 2014a. Species radiation by niche shifts and New Zealand's rockcresses (*Pachycladon*, Brassicaceae). Syst. Biol. 63, 192–202.

Joly, S., Davies, J.T., Archambault, A., Bruneau, A., Derry, A., Kembel, S.W., Peres-Neto, P., Vamosi, J., Wheeler, T.A., 2014b. Ecology in the age of DNA barcoding: the resource, the promise and the challenges ahead. Mol. Ecol. Resour. 14, 221–232.

Kartesz, J.T., 2013. The Biota of North America Program (BONAP). Taxonomic Data Center. <http://www.bonap.net/tdc>. Chapel Hill, N.C.

Kerkhoff, A.J., Moriarty, P.E., Weiser, M.D., 2014. The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. Proc. Natl. Acad. Sci. USA 111, 8125–8130.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., Webb, C.O., 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26, 1463–1464.

Kress, W.J., Erickson, W.J., 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. PLOS ONE, e508.

Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Phylogenet. Evol. 29, 1695–1701.

Lavoie, C., 2013. Biological collections in an ever changing world: Herbaria as tools for biogeographical and environmental studies. Persp. Plant Ecol. Evol. Syst. 15, 68–76.

Lawing, A.M., Matzke, N.J., 2014. Conservation paleobiology needs phylogenetic methods. Ecography 37, 001–014.
Lewis, P.O., 2003. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913925.
Loera, I., Sosa, V., Ickert-Bond, S.M., 2012. Diversification in North American arid lands: niche conservatism, divergence and expansion of habitat explain speciation in genus *Ephedra*. Mol. Phylogenet. Evol. 65, 437–450.
Losos, J.B., 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. Ecol. Lett. 11, 995–1007.
Marshall, S.J., James, T.S., Clarke, G.K.C., 2002. North American ice sheet reconstructions at the Last Glacial Maximum. Quat. Sci. Rev. 21, 175–192.
Matzke, N.J., 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Front. Biogeogr. 4, 242–247.
Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol.. http://dx.doi.org/10.1093/sysbio/syu056.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA, pp. 1–8.
Morrone, J.J., 2014. Parsimony analysis of endemicity (PAE) revisited. J. Biogeogr. 41, 842–854.
Muasya, A.M., Simpson, D.A., Verboom, G.A., Goetghebeur, P., Naczi, R.F.C., Chase, M. W., Smets, E., 2009. Phylogeny of Cyperaceae based on DNA sequence data: current progress and future prospects. Bot. Rev. 75, 2–21.
Orme, D., Greckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., Pearse, W., 2013. caper: Comparative Analysis of Phylogenetic and Evolution in R. R package version 0.5.2. <http://CRAN.R-project.org/package=caper>.
Paton, A.J., Springate, D., Suddee, S., Otieno, D., Grayer, R.J., Harley, M.M., Willis, F., Simmonds, M.S.J., Powell, M.P., Savolainen, V., 2004. Phylogeny and evolution of basils and allies (Ocimeae, Labiatae) based on three plastid DNA regions. Mol. Phylogenet. Evol. 31, 277–299.
Pedro, J., Starr, J.R., Naczi, R.F.C., Roalson, E.H., Waterway, M., Hipp, A., Hahn, M., Leuders, K., Global *Carex* Group, 2014. First Steps Towards a Global *Carex* (Cyperaceae) Phylogeny. Abstract # 663, presented at Botany 2014, Boise, ID.
Philips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175.
Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6, 7–11.
Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS ONE 5, e9490.
Pyke, G.H., Ehrlich, P.R., 2010. Biological collections and ecological/environmental research: a review, some observations and a look to the future. Biol. Rev. 85, 247–266.
Qian, H., Wiens, J.J., Zhang, J., Zhang, Y., 2014. Evolutionary and ecological causes of species richness patterns in North American angiosperm trees. Ecography 37, 001–010.
Quantum GIS Development Team, 2014. Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project, available <http://www.qgis.org>.
de Queiroz, A., Gatesy, J., 2007. The supermatrix approach to systematics. Trends Ecol. Evol. 22, 34–41.
R Development Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rabosky, D.L., Santini, F., Eastman, J., Smith, S.A., Sidlauskas, B., Chang, J., Alfaro, M. E., 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat. Commun. http://dx.doi.org/10.1038/ncomms2958.
Rabosky, D.L., Donnellan, S.C., Grundler, M., Lovette, I.J., 2014a. Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. Syst. Biol. 63, 610–627.
Rabosky, D.L., Grundler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W., Huang, H., Larson, J.G., 2014b. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. Metho. Ecol. Evol. 5, 701–707.
Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14.
Revell, L.J., 2012. Phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3, 217–233.
Rocha, L.A., Aleixo, A., Allen, G., Almeda, F., Baldwin, C.C., Barclay, M.V.L., Bates, J.M., Bauer, A.M., Benzoni, F., Berns, C.M., Berumen, M.L., Blackburn, D.C., Blum, S., Bolanos, F., Bowie, R.C.K., Britz, R., Brown, R.M., Cadena, C.D., Carpenter, K., Ceriaco, L.M., Chakrabarty, P., Chaves, G., Choat, J.H., Clements, K.D., Collette, B. B., Collins, A., Coyne, J., Cracraft, J., Daniel, T., de Carvalho, M.R., De Queiroz, K., Di Dario, F., Drewes, R., Dumbacher, J.P., Engilis, A.J., Erdmann, M.V., Eschmeyer, W., Feldman, C.R., Fisher, B.L., Fjeldsa, J., Fritsch, P.W., Fuchs, J., Getahun, A., Gill, A., Gomon, M., Gosliner, T., Graves, G.R., Griswold, C.E., Guralnick, R., Hartel, K., Helgen, K.M., Ho, H., Iskandar, D.T., Iwamoto, T., Jaafar, Z., James, H.F., Johnson, D., Kavanaugh, D., Knowlton, N., Lacey, E., Larson, H.K., Last, P., Leis, J.M., Lessios, H., Liebherr, J., Lowman, M., Mahler, D.L., Mamonekene, V., Matsuura, K., Mayer, G.C., Mays, H.J., McCosker, J., McDiarmid, R.W., McGuire, J., Miller, M.J., Mooi, R., Mooi, R.D., Moritz, C., Myers, P., Nachman, M.W., Nussbaum, R.A., Foighil, D.O., Parenti, L.R., Parham, J.F., Paul, E., Paulay, G., Perez-Eman, J., Perez-Matus, A., Poe, S., Pogonoski, J., Rabosky, D.L., Randall, J.E., Reimer, J.D., Robertson, D.R., Roedel, M.O., Rodrigues, M.T., Roopnarine, P., Rueber, L., Ryan, M.J., Sheldon, F., Shinohara, G., Short, A., Simison, W.B., Smith-Vaniz, W.F., Springer, V.G., Stiassny, M., Tello, J.G., Thompson, C.W., Trnski, T., Tucker, P., Valqui, T., Vecchione, M., Verheyen, E., Wainwright, P.C., Wheeler, T.A., White, T., Will, K., Williams, J.T., Williams, G., Wilson, E.O., Winker, K., Winterbottom, R., Witt, C.C., 2014. Specimen collection: an essential tool. Science 344, 814–815.
Slater, G.J., Price, S.A., Santini, F., Alfaro, M.E., 2010. Diversity versus disparity and the radiation of modern cetaceans. Proc. R. Soc. B 277, 3097–3104.
Smith, S.A., Donoghue, M.J., 2010. Combining historical biogeography with niche modeling in the *Caprifolium* clade of *Lonicera* (Caprifoliaceae, Dipsacales). Syst. Biol. 59, 322–341.
Smith, S.A., O'Meara, B.C., 2012. TreePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28, 2689–2690.
Spalink, D., Drew, B.T., Pace, M.C., Zaborksy, J.G., Cameron, K.C., Starr, J.R., Sytsma, K. J. (submitted for publication). Conquering then dividing: explaining the area-richness correlation in a cosmopolitan family of vascular plants. J. Biogeogr.
Stamatakis, A., 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.
Tang, G., Zhang, M.G., Zhou, Z., Chen, W., Slik, J.W.F., 2014. Phylogenetic support for the tropical niche conservatism hypothesis despite the absence of a clear latitudinal richness gradient in Yunnan's woody flora. Biogeosci. Discuss. 11, 7055–7077.
Tkach, N., Ree, R.H., Kuss, P., Röser, M., Hoffman, M.H., 2014. High mountain origin, phylogenetics, evolution, and niche conservatism of arctic lineages in the hemiparasitic genus *Pedicularis* (Orobanchaceae). Mol. Phylogenet. Evol. 76, 75–92.
Töpel, M., Antonelli, A., Yesson, C., Eriksen, B., 2012. Past climate change and plant evolution in western North America: a case study in Rosaceae. PLoS ONE 7, e50358.
Vamosi, J.C., Vamosi, S.M., 2011. Factors influencing diversity in angiosperms: at the crossroad of intrinsic and extrinsic traits. Am. J. Bot. 98, 460–471.
Viljoen, J., Muasya, M.A., Barrett, R.L., Bruhl, J.J., Gibbs, A.K., Slingsby, J.A., Wilson, K. A., Verboom, G.A., 2013. Radiation and repeated transoceanic dispersal of Schoeneae (Cyperaceae) through the southern hemisphere. Am. J. Bot. 100, 2494–2508.
Villalobos, F., Rangel, T.F., Diniz-Filho, J.A.F., 2013. Phylogenetic fields of species: cross-species patterns of phylogenetic structure and geographical coexistence. Proc. R. Soc. B 280, 20122570.
Villaverde, T., Escudero, M., Martin-Bravo, S., Bruederle, L.P., Luceño, M., Starr, J.R., 2015. Direct long-distance dispersal best explains the bipolar distribution of *Carex arctogena* (*Carex* sect. *Capituligerae*, Cyperaceae). J. Biogeogr. 42, 1514–1525.
Voss, E.G., Reznicek, A.A., 2012. Field Manual of Michigan Flora. University of Michigan Press, Ann Arbor.
Webb, C.O., Ackerly, D.D., McPeek, M.A., Donoghue, M.J., 2002. Phylogenies and Community Ecology. Annu. Rev. Ecol. Syst. 33, 475–505.
White, T.J., Bruns, T., Lee, S., Taylor, J.W., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR Protocols: a Guide to Methods and Applications. Academic Press, New York, pp. 315–322.
Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Syst. Biol. syr025.
Wiens, J.J., Tiu, J., 2012. Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling. PLoS ONE 7, e42925.
Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Davies, T.J., Grytnes, J., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M., Stephens, P.R., 2010. Niche conservatism as an emerging principle in ecology and conservation biology. Ecol. Lett. 13, 1310–1324.
Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686–693.