# Research Article

# A tale of worldwide success: Behind the scenes of *Carex* (Cyperaceae) biogeography and diversification

Santiago Martín-Bravo[1,2†]* [ID], Pedro Jiménez-Mejías[2,3,4†]* [ID], Tamara Villaverde[5] [ID], Marcial Escudero[6] [ID], Marlene Hahn[7], Daniel Spalink[8] [ID], Eric H. Roalson[2] [ID], Andrew L. Hipp[7,9] [ID], and the Global *Carex* Group (Carmen Benítez-Benítez[1], Leo P. Bruederle[10], Elisabeth Fitzek[7,11], Bruce A. Ford[12], Kerry A. Ford[13], Mira Garner[7,14], Sebastian Gebauer[15], Matthias H. Hoffmann[16], Xiao-Feng Jin[17], Isabel Larridon[18,19], Étienne Léveillé-Bourret[20], Yi-Fei Lu[17], Modesto Luceño[1], Enrique Maguilla[6], Jose Ignacio Márquez-Corro[1], Mónica Míguez[1], Robert Naczi[21], Anton A. Reznicek[22], and Julian R. Starr[23])

[1]Área de Botánica, Department of Molecular Biology and Biochemical Engineering, Universidad Pablo de Olavide, Ctra de Utrera km 1 sn, Seville 41013, Spain
[2]School of Biological Sciences, Washington State University, Pullman, WA 99164, USA
[3]Department of Biology (Botany), Universidad Autónoma de Madrid, Campus Cantoblanco, Madrid 28049, Spain
[4]Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad Autónoma de Madrid, Madrid 28049, Spain
[5]Real Jardín Botánico, CSIC, Plaza de Murillo 2, Madrid 28014, Spain
[6]Department of Plant Biology and Ecology, University of Seville, Reina Mercedes sn, Seville ES-41012, Spain
[7]The Morton Arboretum, 4100 Illinois Route 53, Lisle, IL 60532, USA
[8]Department of Ecosystem Science and Management, Texas A&M University, 495 Horticulture Rd Suite 305, College Station, TX 77843, USA
[9]The Field Museum, 1400 S Lake Shore Dr, Chicago, IL 60605, USA
[10]Department of Integrative Biology, University of Colorado Denver, Denver, CO 80217-3364, USA
[11]Universität Bielefeld, Department of Computational Biology & Center for Biotechnology - CeBiTec, Universitaetsstrasse 27, Bielefeld 33615, Germany
[12]Department of Biological Sciences, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada
[13]Allan Herbarium, Manaaki-Whenua Landcare Research, PO Box 69040, Lincoln 7640, Canterbury, New Zealand
[14]Current address: The University of British Columbia, Forest Sciences Centre, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada
[15]Department of Systematic Botany, Martin Luther University Halle-Wittenberg, Geobotany and Botanical Garden, Neuwerk 21, Halle 06108 (Saale), Germany
[16]Martin Luther University Halle-Wittenberg, Geobotany and Botanical Garden, Botanical Garden, Am Kirchtor 3, Halle 06108 (Saale), Germany
[17]College of Life and Environment Sciences, Hangzhou Normal University, Hangzhou 311121, Zhejiang, China
[18]Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, United Kingdom
[19]Department of Biology, Ghent University, Systematic and Evolutionary Botany Lab, K.L. Ledeganckstraat 35, Gent 9000, Belgium
[20]Institute of Systematic and Evolutionary Botany, University of Zürich, Zollikerstrasse 117, Zürich 8008, Switzerland
[21]New York Botanical Garden, 2900 Southern Blvd., Bronx, NY 10458, USA
[22]University of Michigan Herbarium, 3600 Varsity Drive, Ann Arbor, MI 48108-2228, USA
[23]Department of Biology, University of Ottawa, Gendron Hall, Room 160, 30 Marie Curie, Ottawa, Ontario, K1N 6N5, Canada
[†]These authors contributed equally to this work.
*Authors for correspondence. E-mail: smarbra@upo.es (SM-B); E-mail: pjimmej@gmail.com (PJ-M)

**Abstract**   The megadiverse genus *Carex* (c. 2000 species, Cyperaceae) has a nearly cosmopolitan distribution, displaying an inverted latitudinal richness gradient with higher species diversity in cold-temperate areas of the Northern Hemisphere. Despite great expansion in our knowledge of the phylogenetic history of the genus and many molecular studies focusing on the biogeography of particular groups during the last few decades, a global analysis of *Carex* biogeography and diversification is still lacking. For this purpose, we built the hitherto most comprehensive *Carex*-dated phylogeny based on three markers (ETS–ITS–*matK*), using a previous phylogenomic Hyb-Seq framework, and a sampling of two-thirds of its species and all recognized sections. Ancestral area reconstruction, biogeographic stochastic mapping, and diversification rate analyses were conducted to elucidate macroevolutionary biogeographic and diversification patterns. Our results reveal that *Carex* originated in the late Eocene in E Asia, where it probably remained until the synchronous diversification of its main subgeneric lineages during the late Oligocene. E Asia is supported as the cradle of *Carex* diversification, as well as a "museum" of extant species diversity. Subsequent "out-of-Asia" colonization patterns feature multiple asymmetric dispersals clustered toward present times among the Northern Hemisphere regions, with major regions acting both as source and sink (especially Asia and North America), as well as several independent colonization events of the Southern Hemisphere. We detected 13 notable diversification rate shifts during the last 10 My,

including remarkable radiations in North America and New Zealand, which occurred concurrently with the late Neogene global cooling, which suggests that diversification involved the colonization of new areas and expansion into novel areas of niche space.

**Key words:** ancestral area reconstruction, biogeographic stochastic mapping, boreo-temperate, dispersal, diversification rates, hyperdiverse, phylogeny.

# 1 Introduction

The sedge genus *Carex* L. (Cyperaceae), with nearly 2000 species, is among the three largest angiosperm genera in the world (POWO, 2019; WCSP, 2019). Remarkably, it has undergone few taxonomic rearrangements since its formal description by Linnaeus in 1753. Beyond a few minor satellite genera recently subsumed into *Carex* (c. 150 spp.; *Cymophyllus* Mack., *Kobresia* Willd., *Schoenoxiphium* Nees, and *Uncinia* Pers.; Global Carex Group-GCG herein-, 2015), the morphological homogeneity of the diagnostic characteristics of *Carex* (unisexual flowers, with the pistillate ones enclosed in a bract-derived structure called the perigynium or utricle; Kükenthal, 1909, Egorova, 1999, see also Jiménez-Mejías et al., 2016a) have reinforced the integrity of the genus and prevented its decomposition into a number of smaller genera. *Carex* represents a remarkable success as an evolutionary model, maintaining morphological coherence as a genus while diversifying ecologically and taxonomically.

*Carex*, as currently circumscribed (GCG, 2015), is a nearly cosmopolitan genus. It is most diverse in the Northern Hemisphere boreo-temperate zone and, to a lesser extent, the Southern Hemisphere temperate zone. As such, it exhibits an inverted latitudinal species richness gradient, with higher diversity in cold-temperate areas and tropical regions harboring few species (Kindlmann et al., 2007; Escudero et al., 2012; see Fig. 1). Approximately 20 years of phylogenetic studies at multiple sampling scales, from species complexes to the genus as a whole (Starr et al., 1999, 2004, 2008, 2015; Roalson et al., 2001; Hendrichs et al., 2004a, 2004b; Waterway & Starr, 2007; Derieg et al., 2008; Dragon & Barrington, 2009; Escudero & Luceño, 2009; Gehrke & Linder, 2009; Waterway et al., 2009; Gehrke et al., 2010; Ford et al., 2012; Jiménez-Mejías et al., 2012b; Martín-Bravo et al., 2013; Gebauer et al. 2014; Yano et al., 2014; Gebauer et al., 2015; Maguilla et al., 2015; Villaverde et al., 2015a, 2015b, 2017b, 2017c, in review; GCG, 2016a, 2016b; Míguez et al., 2017), have predominantly focused on investigating systematic relationships, morphological evolu-tion, and biogeography at fine phylogenetic scales. These studies have demonstrated that while historical sectional classification units were often geographically delimited, shallow clades within the genus may sometimes be found distributed on multiple continents. This, together with the frequent homoplasy observed for certain characteristics, led to the creation of artificial sections grouping apparently similar species that displayed more or less congruent distributions and/or ecologies. As a direct consequence of this, historical sectional classifications are largely incongruent with what we know today about phylogenetic relationships in *Carex* (see GCG, 2016a).

Biogeography appears to have played a critical role in the diversification of *Carex*. The origin and worldwide diversification of *Carex* have been discussed in relation to chromosome evolution and adaptation to colder climates (Escudero et al., 2012; Gebauer et al., 2014; Hoffmann et al., 2017), as well as fine-scale partitioning of niche and distributional ranges of individual species within continents (Waterway et al., 2009; Gebauer et al., 2014; Pender, 2016; Spalink et al., 2016a, 2016b, 2018; Benítez-Benítez et al., 2018). While Cyperaceae have been inferred to have arisen in South America at the late Cretaceous, it was the migration of the ancestors of *Carex* to the Northern Hemisphere that was the catalyst for the major diversification of this lineage (Léveillé-Bourret et al., 2014, 2015, 2018a, 2018c; Spalink et al., 2016b). Recent phylogenetic studies have discovered that several early-diverged lineages of *Carex* and its closest living relatives are Southeast (SE) Asian (Starr & Ford, 2009; Waterway et al., 2009; Starr et al., 2015; Léveillé-Bourret et al., 2018b), supporting the long-held view that the genus originated in SE Asia (Nelmes, 1951; Raymond, 1955, 1959; Koyama, 1957; Ball, 1990). A diversification shift has been documented near the crown of the genus (Escudero et al., 2012; Spalink et al., 2016b; Márquez-Corro et al., 2019), which is possibly associated with its transition into cooler climates and a shift to high rates of fissions and fusions as the dominant mode of chromosome evolution (Hipp et al., 2009; Escudero et al., 2012; Márquez-Corro et al., 2019). However, this origin story, which has become well understood over the past decade, does not explain the astonishing biogeographic diversity of the genus. *Carex* displays many of the large-scale diversity patterns observed in angiosperms, with different groups showing various diversity gradients, centers of endemism, and diversity hotspots. Despite the polyphyly of many *Carex* sections (GCG, 2016a), species relationships across geographic areas have in many cases been predicted fairly accurately by morphological taxonomy before the development and popularization of molecular systematic techniques (see Reznicek, 1990; Egorova, 1999; Ball & Reznicek, 2002). Having a striking capacity for long-distance dispersal (Villaverde et al., 2017a), *Carex* has colonized the Southern Hemisphere several times from different source areas (Spalink et al., 2016b). Gehrke & Linder (2009), for example, showed that all northern continents were probably involved in the colonization of sub-Saharan Africa. Escudero et al. (2009) demonstrated dispersal within sect. *Spirostachyae* from W Palearctic to tropical Africa and South America. Míguez et al. (2017) demonstrated transitions in sect. *Rhynchocystis* between Europe and Tropical Africa and Jiménez-Mejías et al. (2012b) between the Northern Hemisphere, South America, South Africa, and New Zealand in sect. *Ceratocystis*. Bipolar distributions at the species level have also been particularly well-documented in *Carex* and explained mostly by direct long-distance dispersal from the Northern Hemisphere to high latitudes of the Southern Hemisphere in South America and New Zealand (Villaverde et al., 2015a, 2015b, 2017b, 2017c; Márquez-Corro et al., 2017; Maguilla et al., 2018). The genus also exhibits circumpolar (Gebauer et al., 2014; Hoffmann et al., 2017; Maguilla et al., 2018), Beringian (Schönswetter et al., 2008; King & Roalson, 2009; Maguilla et al., 2018; amphi-Atlantic (Schönswetter et al., 2008; Jiménez-Mejías et al., 2012b; Westergaard et al., 2019), Arctic-Alpine (Schöns-wetter et al., 2006, 2008; Jiménez-Mejías et al., 2012a; Gebauer

et al., 2014; Hoffmann et al., 2017), pan-Himalayan (Uzma et al., 2019), Europe-Central Asia (Schönswetter et al., 2006), East-West Europe/Mediterranean (Escudero et al., 2009, 2010; Jiménez-Mejías et al., 2011, 2012a; Míguez et al., 2017; Benítez-Benítez et al., 2017, 2018), and Eastern-Western North America (Roalson & Friar, 2004a, 2004b; Hipp et al., 2006; Hipp, 2008; Dragon & Barrington, 2009) distribution patterns, all illuminated using phylogenetic approaches. Colonization of isolated oceanic archipelagos from mainland sources has also been documented by several authors; these include Hawaii (Dragon & Barrington, 2009), Macaronesia (Escudero et al., 2009; Jiménez-Mejías et al.,

2012b; Míguez et al., 2017), Mascarenes and Tristan da Cunha (Escudero et al., 2009), and Juan Fernández (Ridley & Jiménez-Mejías, in prep.). The genus is a treasure trove of biogeographic scenarios. Yet, many striking disjunctions identified by previous authors and apparently supported by the most recent phylogenies (e.g., GCG, 2016a) are not well understood, such as the Gondwanan, circumantarctic, pantropical, trans-Pacific, Madrean-Tethyan, trans-Caribbean, or Western-Eastern Eurasian distributions found in several *Carex* species groups and clades. The recent inference of a phylogeny comprising about 50% of all the accepted species (GCG, 2016a) and a specimen-level
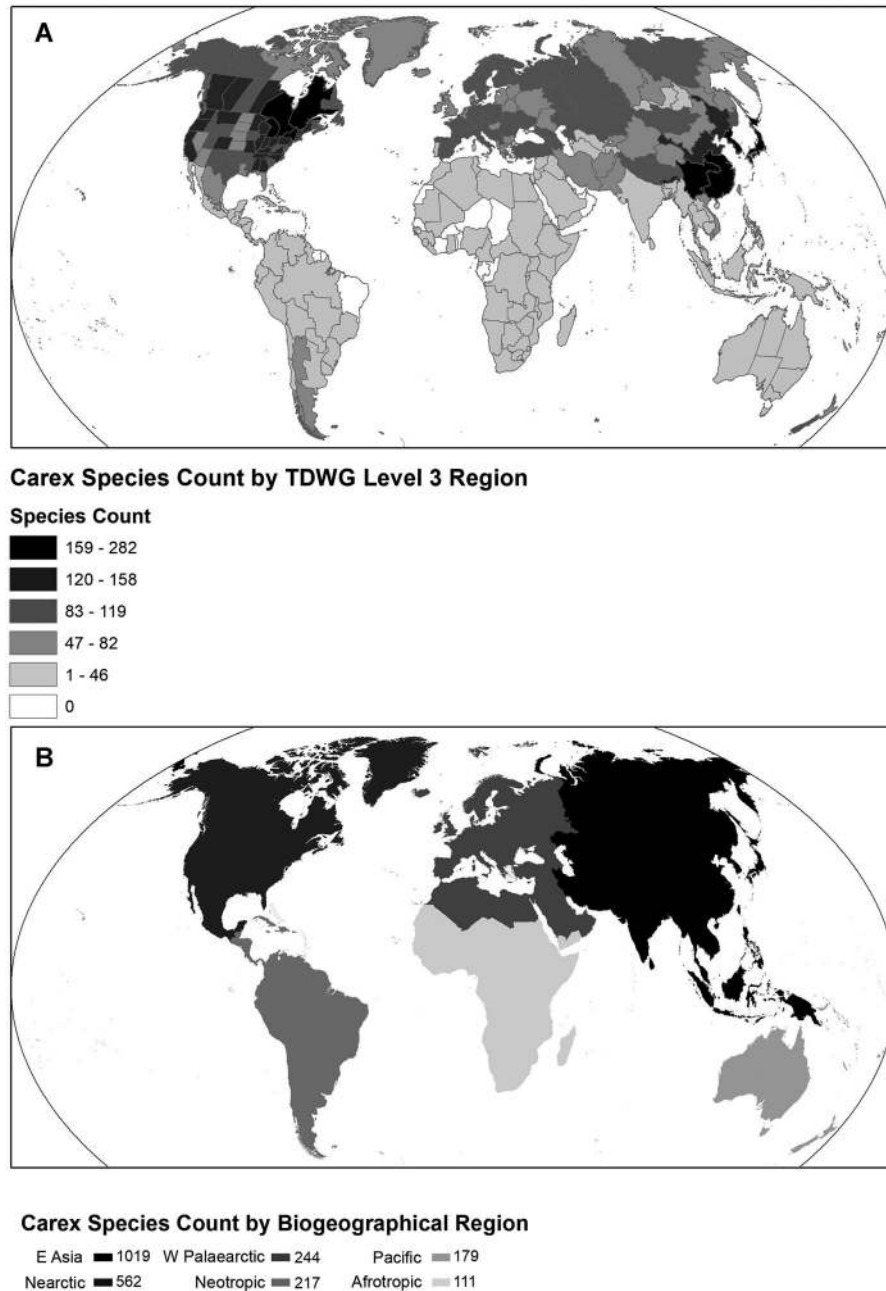


**Fig. 1.** Global diversity of *Carex* species represented according to **A,** TDWG level 3 regions ("botanical countries"; Brummitt, 2001); and **B,** Biogeographical regions as coded in the present work. Note in **A,** the much higher species diversity in cold-temperate zones of both Hemispheres (especially the Northern Hemisphere) and the much lower diversity of tropical areas.

aggregation of the vast majority of the previously published sequences available in NCBI GenBank comprising about 800 taxa (GCG, 2016b) provide the opportunity to explore these biogeographic scenarios more broadly. At the same time, the development of novel phylogenomic tools has given us a robust understanding of the framework phylogenetic structure of the genus (Villaverde et al., in review). While most relationships at deep nodes remain poorly resolved, with significant conflicts among studies (see Villaverde et al., in review), the newly available data resulted in a robust phylogenetic framework that provides a significant clarification of the relationships among the main *Carex* lineages, also serving as the base for a revised subgeneric classification (Villaverde et al., in review).

In this study, we increase the sampling of the global phylogeny of the GCG (2016a, 2016b) by more than 300 species and constrain the backbone of the phylogeny using this phylogenomic framework (Villaverde et al., in review) to study broad-scale patterns of historical biogeography and diversification. The result is the most comprehensively sampled time-calibrated phylogenetic hypothesis of the world's *Carex*, which we expect to serve as the basis for subsequent analyses of the diversification of the genus.

# 2 Material and Methods

## 2.1 Study group nomenclature
We follow the most recent subgeneric rearrangement proposed by Villaverde et al. (in review), which treats each of the six main *Carex* clades as different subgenera: subg. *Siderosticta* Waterway, subg. *Carex*, subg. *Euthyceras* Peterm., subg. *Psyllophorae* (Degl.) Peterm., subg. *Uncinia* (Pers.) Peterm., and subg. *Vignea* (P. Beauv. ex T. Lestib.) Heer. Accordingly, we will refer to these main clades directly using the names of the subgenera.

Species groups are named using *Carex* sectional names according to the sectional partition provided in Data S1, which is a modified version of the one presented in Global Carex Group (2016a). Given the intricate nomenclature of *Carex* sections, and for the sake of clarity, the use of sectional authorities is avoided, as the current study is not intended to present nomenclatural rearrangements. Since a number of sections are polyphyletic, and in order to simplify the text, whenever a clade contains the vast majority of the species belonging to a single section, we will refer to the clade as that section. When further explanation is needed to specify what subset of species we are talking about, we will resort to commonly used terms for systematic grouping: "alliances" for monophyletic groups of intermingled sections, "core" for the clade containing the type species of a nonmonophyletic section, or geographic references (e.g., "North American clade of section *x*").

## 2.2 Sampling
We built a comprehensive *Carex* phylogeny based on the DNA barcoding nrDNA ITS and ETS and cpDNA *matK*, using the dataset published by GCG (2016a) as the starting point. This initial matrix consisted of 2150 concatenated *Carex* sequences, representing 996 of the currently 1992 accepted species (50%; Govaerts et al., 2019) from 110 of the 126 recognized sections (92.06%; see GCG, 2016a). We expanded

this dataset by adding 2402 concatenated sequences available in GenBank for these DNA regions up to mid-2014 as compiled in GCG (2016b); 522 new concatenated sequences (see Data S2) obtained for this study from materials on loan from A, BISH, E, K, MO, NY, and TUS (Thiers, 2019), following the lab protocols described in GCG (2016a); and 448 concatenated sequences available in GenBank for these DNA regions (ITS, ETS, *matK*) and published in recent phylogenetic studies (Léveillé-Bourret et al., 2014, 2018a, 2018c; Gebauer et al., 2015; Molina et al., 2015; Starr et al., 2015; Villaverde et al., 2015a, 2015b, 2017c; Elliott et al., 2016; Benítez-Benítez et al., 2017; Márquez-Corro et al., 2017; Míguez et al., 2017; see Data S2). The outgroup was composed of seven samples, representing tribe Scirpeae (*Eriophorum vaginatum*, *Scirpus polystachyus*), and the recently described tribes Trichophoreae (*Trichophorum alpinum*, *T. cespitosum*) and Sumatroscirpeae (*Sumatroscirpus paniculatocorymbosus*, *S. rupestris*), which are successive sister clades to *Carex* (Léveillé-Bourret et al., 2018a, 2018c; Léveillé-Bourret & Starr, 2019; Semmouri et al., 2019). This sequence compilation yielded a raw multiaccession matrix (hereafter the "multiple tips matrix") of 5529 individuals comprising one or more concatenated ETS, ITS, and *matK* sequences each.

Due to the huge size of the matrix, this was subjected to several rounds of curation to detect conflicting phylogenetic placements due to contamination, mislabeling, duplication, misidentification, and misconcatenation. We also removed those accessions whose placement in the trees was odd and poorly supported due to low phylogenetic signal (usually when *matK* was the only sequence available). Accessions from the former tribe Cariceae genera (*Cymophyllus*, *Kobresia*, *Schoenoxiphium*, and *Uncinia*) were renamed to conform to the current *Carex* taxonomy (GCG, 2015). Through this process, we selected a single accession for each taxon (species, subspecies or variety), prioritizing individuals with the longest and highest number of sequences. The resulting one-tip-per-taxon matrix (hereafter, the "singletons matrix") was subsequently used for the biogeographic and diversification analyses. The final curated multiple tips matrix was composed of 4467 concatenated sequences, of which 1392 were retained in the singletons matrix, representing 1386 *Carex* taxa (six accessions corresponded to the outgroup) and 1312 *Carex* species belonging to 126 sections. This dataset comprises the first complete sampling of sections for any *Carex* phylogenetic analysis and a 32% increase in species sampling from the GCG (2016a) dataset, almost reaching 66% of the total number of accepted *Carex* species (Govaerts et al., 2019). The sampling proportion varied significantly between the different regions considered in the biogeographic analyses (see results).

## 2.3 Alignment
Owing to the large nucleotide variability across the multiple tips matrix, especially among the sequences of the nrDNA regions (ETS and ITS), and the consequent difficulty of aligning homologous positions, two different alignment approaches were conducted. Each DNA region (ETS, ITS and *matK*) was aligned separately by means of (i) a default alignment using MUSCLE v.3.8.31 (Edgar, 2004) and (ii) following the approach taken in GCG (2016a), which

consisted of a profile-to-profile approach in which the three gene regions were separated into three submatrices corresponding to the major clades found in previous works, aligned with MUSCLE, then aligned to one another using profile-to-profile alignment in MUSCLE, retaining alignments within groups but inserting gaps to align between groups. The matrices and phylogenetic reconstructions (see below) resulting from each of these approaches were visually examined and compared, clearly revealing that the default MUSCLE alignment, our first approach, produced a better-resolved topology (profile-to-profile alignment results not shown).

To build the singletons matrix, the selected sequences were realigned from scratch using the default MUSCLE alignment. The lower proportion of missing data in the singletons matrix compared to the multiple tips matrix considerably reduced the ambiguity of the alignment and improved topology resolution in the phylogenetic analyses (see below).

## 2.4 Phylogenetic analyses

The aligned multiple tips matrix (3006 bp) contained a high amount of missing data (60.08%), since one or two of the three selected DNA regions were frequently missing, either because we failed in the amplification from herbarium specimens or because the selected GenBank accessions only represented one or two of the selected regions. This produced poorly resolved topologies when the matrix was directly analyzed (results not shown). Therefore, we resorted to two different scaffolding approaches (Hinchliff & Roalson, 2013; GCG, 2016a).

For the first, we followed the GCG (2016a) scaffolding approach in which we first selected only those accessions of the multiple tips matrix represented for both ITS and ETS (regardless of their having *matK* or not), to build a nrDNA-complete matrix (2394 accessions, 3006 aligned bp) with a reduced amount of missing data (46.17%). This multiple tips nrDNA-complete matrix was subsequently used to build a backbone tree (reference tree) with maximum likelihood (ML), as implemented in RAxML v. 8.2.10 (Stamatakis, 2014), through the CIPRES Science Gateway (Miller et al., 2010). Then, the phylogenetic placement of all excluded sequences—those present in the complete multiple tips matrix, but not in the multiple tips nrDNA-complete matrix—was obtained building a "query tree" based on the reference tree, by using the evolutionary placement algorithm (Berger et al., 2011), as implemented in RAxML. Finally, the SH-aLRT value (nonparametric Shimodaira-Hasegawa implementation of the approximate likelihood-ratio test; Anisimova & Gascuel, 2006; Anisimova et al., 2011) was used to evaluate query tree branch support with RAxML. This analytical procedure was performed to build two query trees, one from the multiple tips matrix (multiple tips query tree) and another from the singletons matrix (singletons query tree). The singletons matrix (2797 aligned bp) had 41% missing data, while this percentage decreased to 33.38% in the singletons nrDNA-complete matrix (1109 concatenated sequences, 2797 aligned bp).

For the second approach, we constrained ML analyses in RAxML using a backbone topology for the genus *Carex* obtained using a phylogenomic approach (Hyb-Seq; Villa-verde et al., in review). This phylogeny has uncovered a novel and strongly supported backbone topology for *Carex*, demonstrating subg. *Uncinia* to be sister to subg. *Vignea* instead of being nested within subg. *Euthyceras*, as has been suggested by most Sanger-based phylogenies. After constraining the singletons nrDNA-complete matrix using the Hyb-Seq tree, we then used, in turn, the resulting constrained singletons tree as a constraint for the complete singletons matrix to obtain the final constrained singletons tree. Clade bootstrap support and search for the best-scoring ML tree were jointly obtained through 100 fast bootstrap replicates in the same single run. Throughout the remainder of the paper, "constrained singletons tree" refers to this twice-constrained singletons tree.

## 2.5 Dating and diversification analyses

We fossil-calibrated the singletons query tree and the constrained singletons tree (1386 *Carex* taxa, 1312 species) using the recently reassessed *Carex* fossil record (Jiménez-Mejías et al., 2016b) to establish calibration points. Ten fossil constraints were applied (Table 1), with ages ranging from the Eocene for the crown node of *Carex* (*C. colwellensis*), to the Pliocene for sections *Ammoglochin* and *Ovales* (*C. ungeri* and *C. klarae*, respectively). Three fossils were used to calibrate deep nodes (crown node of genus *Carex*, subg. *Carex* and subg. *Vignea*), and seven for shallower nodes (Table 1). The taxonomic identity of these fossils has been recently evaluated (Jiménez-Mejías & Martinetto, 2013; Jiménez-Mejías et al., 2016b). They consist mostly of fossil nutlets, though some preserved utricles were utilized as well (i.e., *C. hartauensis* and *C. flagellata*).

The joint use of fossils at both deep and shallow nodes may produce interaction among the calibrations because the node ages in a tree are not mutually independent (Ho & Phillips, 2009). To assess the sensitivity of our inferences to this potential interaction, we analyzed the trees by using (i) all 10 fossils and (ii) only the three oldest fossils that constrain the deepest nodes (i.e., *C. colwellensis*, *C. marchica*, *C. hartauensis*; Table 1).

Preliminary analyses using Bayesian MCMC in BEAST v. 1.8.4 (Drummond & Rambaut, 2007; Drummond et al., 2012) yielded problems with mixing of chains and low ESS values. Consequently, the analyses presented here utilize the penalized likelihood approach (Sanderson, 2002) as implemented in TreePL (Smith & O´Meara, 2012), which was designed for large phylogenies. The rate smoothing parameter was set on the basis of cross-validation and the $\chi^2$ test in TreePL. Nine smoothing values between $1e^{-5}$ and $1e^3$ were compared.

Transitions in lineage diversification rates were estimated on the constrained singletons tree calibrated with 10 fossils using the speciation-extinction model implemented in Bayesian analysis of macroevolutionary mixtures (BAMM; Rabosky, 2014). The method models transitions in net diversification rates by allowing changes in the numbers and locations of nodes at which speciation and extinction rates shift, averaging over models and parameters using reversible-jump Markov chain Monte Carlo (rjMCMC; Green, 1995). The R package "BAMMtools" (Rabosky et al., 2014) was used for configuration and analysis of MCMC. All priors were set as recommended using the setBAMMpriors

**Table 1** Fossil calibrations used in the dating analysis

| Fossil | Age (Mya) | Placement |
|---|---|---|
| *Carex colwellensis* Chandler* | Eocene (Priabonian; 38.0–33.9) | Crown node of genus *Carex* |
| *Carex marchica* Mai* | Early Miocene (23.0–16.0) | Crown node of subg. *Vignea* |
| *Carex ungeri* Mai & H. Walther | Pliocene (5.3–2.6) | Crown node of sect. *Ammoglochin* (including *C. arenaria* and excluding *C. remota*) |
| *Carex klarae* Mai | Pliocene (5.3–2.6) | Stem node of sect. *Ovales* (excluding its sister clade containing *C. bonplandii*) |
| *Carex hartauensis* Mai* | Late Oligocene (Chattian, 28.1–23.0) | Crown node of subg. *Carex* (not considering the position of *C. bostrychostigma* and *C. dissitiflora*) |
| *Carex praehirta* Mai & Walther | Late Miocene (Messinian, 7.3–5.3) | Stem node of sect. *Paniceae* (excluding its sister clade containing *C. pilosa* and *C. auriculata*) |
| *Carex klettvicensis* Mai | Early Miocene (23.0–16.0) | Crown node of sect. *Phacocystis* (excluding its sister clade containing sects. *Fecundae* and *Limosae*) |
| *Carex* sect. *Rhomboidales* | Late Miocene (Tortonian, 11.6–7.2) | Crown node of the clade containing the vast majority of sects. *Rhomboidales* and *Mitratae* |
| *Carex plicata* Łańcucka-Środoniowa | Late early Miocene (Burdigalian, 20.4–16.0) | Crown node of sect. *Rhynchocystis* |
| *Carex flagellata* C. Reid & E.M. Reid | Early Miocene (23.0–16.0) | Crown node of sect. *Vesicariae-Paludosae* alliance clade (including *C. grayi* and *C. lasiocarpa*) |

Fossil ages according to Jiménez-Mejías et al. (2016b). Placement is described according to the topology of the constrained singletons tree (Figs. 2A, 3A; Data S5); asterisks indicate the fossils constraining deep nodes in the alternative dating approach using only three fossils; Mya, million years ago.

function on the 10-fossils tree, with the exception that analyses were conducted using a prior of one shift in diversification rates and a prior of 50 shifts, to see whether model-averaged rates were affected. Analyses were conducted assuming a global sampling fraction of 0.693 to account for missing taxa. Missing taxa were assumed missing at random from the tree, which may bias our results toward lower estimates of diversification of predominantly E Asian and Neotropical clades, which are relatively undersampled. However, given the difficulty of assigning missing taxa to particular clades in our study, and the fact that biogeographic transitions are very common in *Carex* (see results), it was not practical to estimate clade-level sampling partitions. In addition, it has been found that even in cases where the random sampling assumption is dramatically violated (Hipp et al., 2019) and taxon sampling is near 60%, global sampling partitions can yield results not significantly different from clade-specific sampling partitions. The rjMCMC was run using Metropolis-coupling with four chains of 20 000 000 generations each, saving trees every 20 000 generations for analysis. The rjMCMC results were investigated using effective sample size for log-likelihood and number of model shifts in the R packages "coda" (Plummer et al., 2006) and BAMMtools.

**2.6 Biogeographic analyses**
Ancestral area reconstruction (AAR) was conducted using the R package "BioGeoBEARS" (Matzke, 2013, 2014a). Due to the nearly cosmopolitan distribution and the wide diversity of biogeographic patterns in *Carex*, coding of areas for the biogeographic analyses to accurately represent the spatial distribution and endemicity of the genus was problematic. An initial coding with 10 areas (Nearctic, W Palearctic, E Palearctic, Afrotropic, Tropical Asia, Polynesia, Neotropic, Australia, New

Zealand, and Subantarctic Islands) was computationally too demanding. Therefore, we merged them, resulting in six areas: Nearctic, W Palearctic, E Asia (including E Palearctic and Tropical Asia), Afrotropic, Neotropic (including Central America and the Caribbean), and Pacific (including Australia, New Zealand, and Polynesia). For the most part, these areas correspond to classic biogeographic realms (Takhtajan, 1986), except for the treatment of E Asia and Pacific regions. They also correspond with major diversity patterns in *Carex*: Australia and New Zealand share many species groups, and a taxonomic turnover in Eurasia is observed between the W and the E Palearctic (see species listed in Chater, 1980; Egorova, 1999; Dai et al., 2010). We did not consider any adjacency coding given the ability of *Carex* to disperse long distances, involving numerous transoceanic dispersals and sometimes between nonadjacent landmasses (as reviewed in the Introduction). All of the sampled taxa were coded as present or absent in these areas (Data S1), based on distributions in Govaerts et al. (2019). We analyzed our dataset under the DIVA-like (dispersal-vicariance analysis; Ronquist, 1997) and DEC models (dispersal-extinction-cladogenesis; Ree et al., 2005; Ree & Smith, 2008) as well as in combination with the "jump dispersal" or founder (j) parameter (models: DEC, DEC+J, DIVALIKE, DIVALIKE+J; Matzke, 2014b). The estimated likelihood of models DEC and DIVALIKE were not compared with estimated likelihood of models DEC+J and DIVALIKE+J, as their likelihoods are not directly comparable based on the way that the J vs. D and E parameters enter into the model (Ree & Sanmartín, 2018). In addition, results from the models considering the parameter J will be interpreted cautiously as they have a greater tendency toward explaining the data entirely by cladogenetic events and inferring, in some cases, anagenetic rates of 0 (Ree & Sanmartín, 2018). Following these analyses, we performed biogeographic stochastic mapping

(BSM) based on the DEC and DEC+J model parameters (Dupin et al., 2016; Matzke, 2016) to estimate the frequency, timing, and locations of anagenetic and cladogenetic events. For each model, we simulated 100 possible histories, given the phylogeny and model parameters, under the constraint that these histories result in the same distributions that we input for the tips of the phylogeny. From these stochastic maps, we calculated the relative frequency of anagenetic and cladogenetic dispersal, sympatry, and vicariance events at 1 million-year interval. We also calculated the frequency of biogeographic regions that were either the source or destination of anagenetic and cladogenetic dispersal events through time.

# 3 Results

The results presented here are based on a dataset including sequences for 65.8% of all recognized *Carex* species, which is, to date, the largest dataset compiled for the genus. Though 34.2% of the species remain to be sampled (mainly Neo-tropical and E Asian taxa), this dataset includes taxa belonging to all accepted sections, suggesting that the phylogenetic and morphological diversity of the genus is relatively well covered. Although we cannot rule out the possibility that the addition of more taxa may reveal lineages that could be sister to one or several of the six currently known subgeneric lineages, we expect that such new lineages would not significantly alter the currently established patterns of relationship.

## 3.1 Phylogenetic and divergence-time analyses
The multiple tips query tree built from the multiple tips matrix (Data S3) and the singletons query tree (Data S4) built using the GCG (2016a) scaffolding approach revealed topologies broadly concordant with the GCG (2016a) tree with respect to the strongly supported monophyly of *Carex* and the five major lineages found to comprise the genus in previous Sanger-based phylogenies (matching subg. *Carex*, *Euthyceras* (with *Uncinia* nested within), *Psyllophorae*, *Siderosticta*, and *Vignea*). Species groupings, shallow clades and unresolved relationships among the five major clades (except for the well-known sister relationship of subg. *Siderosticta* to the rest of *Carex*) were also mostly congruent with GCG (2016a). On the other hand, the constrained singletons tree (Fig. 2A; Data S5) recovered all except three of 1312 sampled species placed within each of the six main subgeneric clades retrieved by Villaverde's et al. (in review) tree, even though the phylogenomic tree was built using only 88 *Carex* species. The only exceptions were the orphan species *C. bostrychostigma*, *C. dissitiflora*, and *C. satsumensis*, which remained unresolved. These species arose from deep nodes in the constrained singletons tree, not forming part of any of the subgeneric lineages (Fig. 2A; Data S5), whereas they appeared at the base of the subg. *Carex* clade in the query trees, but with low support (Data S3 and S4). At the shallowest levels, the query trees tended to be better resolved than the constrained singletons tree, although the groups recovered in the constrained singletons tree were also shared with the query trees and made taxonomic sense.

Inferred diversification times were more influenced by the calibration strategy employed in the dating analyses (10 vs. 3

fossils; Table 1) than by the topology of the dated tree (singletons query tree vs. constrained singletons tree). Thus, the ages recovered under each strategy on each of the two different dated trees were mostly similar, except for differences attributable to the species composition of the dated lineage (Table 2; Data S6–S9). The ages obtained by calibrating the tree using all 10 fossils were considerably older than those obtained using only three fossils at deeper nodes in the phylogeny.

For the remainder of this paper, we will report results based on the most informative dataset we considered: the 10-fossil calibration of the twice-constrained singletons tree (Fig. 2A; Data S6), which takes advantage of both phylogenomic data (Villaverde et al., in review) and a thorough review of the fossil record (Jiménez-Mejías et al., 2016b). In this analysis, the crown node of *Carex* was placed at the late Eocene (Priabonian, 37.17 Mya). Remarkably, the crown nodes of the six major subgeneric lineages within *Carex* were all dated to the late Oligocene (Chattian)-Early Miocene (Aquitanian), with a mean age 22.91–25.18 Mya. These ages are generally older than those obtained in previous studies dating *Carex* (see Table 2).

## 3.2 Biogeographic analyses
Ancestral area reconstruction with DEC (Figs. 2A, S1), DEC+J (Fig. S2), DIVA-like (Fig. S3), and DIVA-like+J (Fig. S4) models produced broadly congruent estimations, though multiple areas were inferred as ancestral more commonly under DIVA-like models than under the DEC models. No significant differences in the most likely estimated area were found among nodes of interest. DEC models had a better fit (DEC, DEC+J: LnL = −2271.46, −2177.61; Akaike information criterion [AIC] = 4546.92, 4361.24; respectively) than DIVA-like models (DIVA-like, DIVA-like+J: LnL = −2363.74, −2255.76; AIC = 4731.48, 4517.52; respectively) so we will hereafter report the results from DEC analyses.

The origin of *Carex* is placed in E Asia, as are the ancestral nodes of subg. *Carex*, *Euthyceras*, *Siderosticta*, and *Vignea* clades (Figs. 2A, S1, S2). E Asia is also maintained as the ancestral area through a number of lineages along all the major clades. Indeed, only subg. *Psyllophorae* and the vast majority of species in subg. *Uncinia* are currently absent from that area. The ancestral distribution of subg. *Uncinia* is inferred to be the Americas by DEC (Figs. 2A, S1), with the most probable area being both the Nearctic–Neotropic and the second most probable area the Neotropic. On the contrary, DEC+J inferred the Neotropic as the most probable ancestral area (Fig. S2). Subgenus *Psyllophorae* is inferred to have arisen in the W Palearctic (Figs. 2A, S1, S2). A number of larger radiations (>15–20 species) have taken place within the last 10 Mya, mostly involving a single geographical area. These areas include E Asia (core sect. *Clandestinae*, core sect. *Mitratae*, core *Kobresia*), the Nearctic (e.g., American sect. *Acrocystis*, sects. *Griseae-Granulares-Careyanae* alliance, sects. *Porocystis-Hymenochlaenae-Longicaules* alliance, sect. *Ovales*), and to a lesser extent the Pacific (specifically two synchronic radiations in New Zealand: sects. *Echinochlaenae* and *Uncinia*) and the Afrotropic (sect. *Indicae* p.p., sect. *Schoenoxiphium*). Such large radiations seem to be conspicuously absent from the W Palearctic and the Neotropic (Figs. S1, S2).
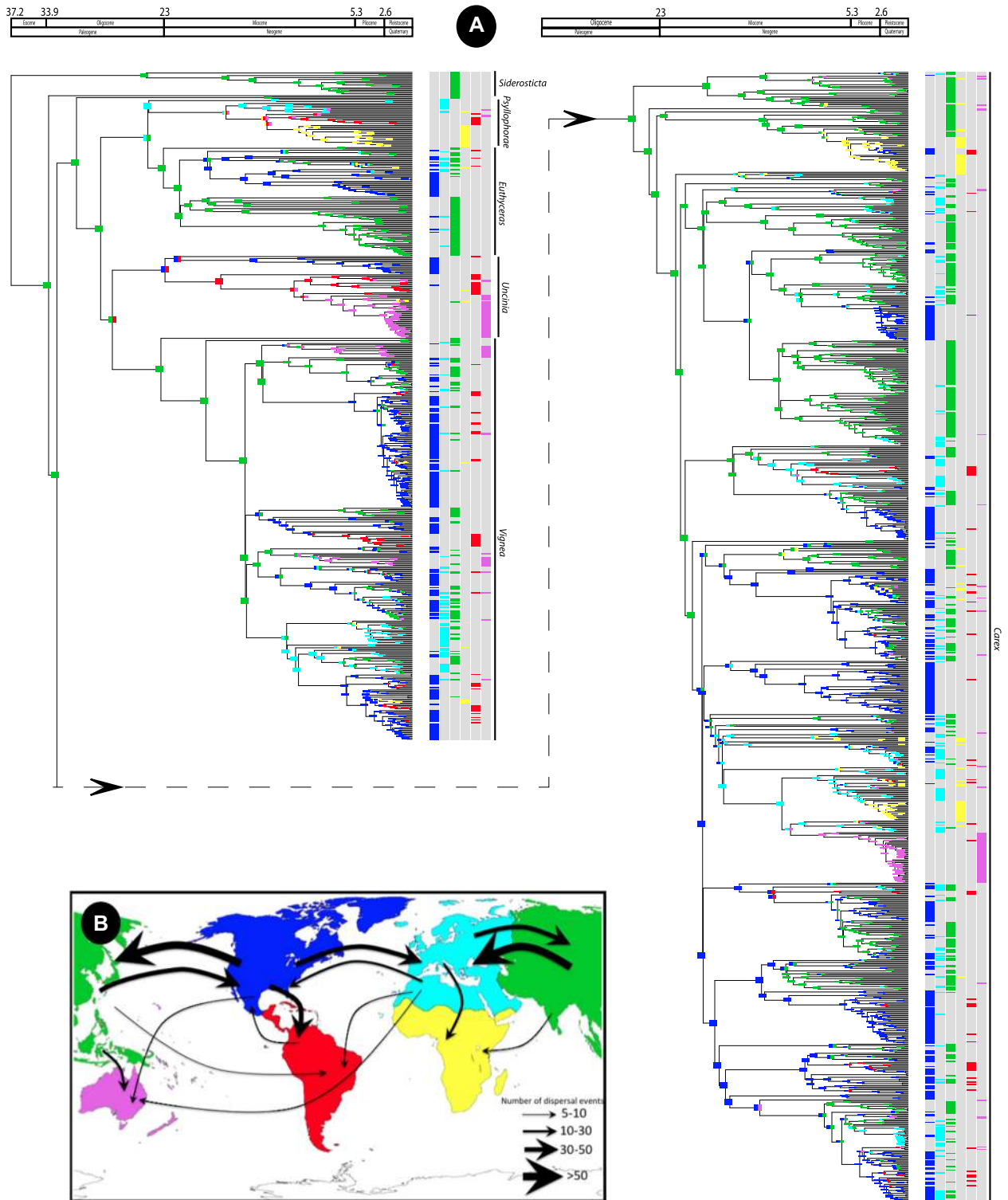
**Fig. 2. A,** Chronogram based on the ML constrained singletons phylogenetic tree of the genus *Carex* calibrated with 10 fossils (Data S6). Geological time scale is displayed at the top of the figure. Subgenera are presented on the vertical black bar at the right. Colored squares at nodes represent the ancestral areas as inferred by DEC model showing the most probable area or combination of areas on each node (Fig. S1; colors according to map in **B**). For branch support see Data S5. **B,** Map displaying the regions coded for the biogeographic analyses in different colors (dark blue: Nearctic; light blue: W Palearctic; green: E Asia; yellow: Afrotropic; red: Neotropic; lilac: Pacific). Arrows represent the dispersal events among the different regions as inferred by the BSM analyses. Arrow thickness is proportional to the number of inferred dispersal events, as shown in the legend. Connections between regions with less than five inferred dispersal events are not illustrated for clarity. BSM, biogeographic stochastic mapping; DEC, dispersal-extinction-cladogenesis; ML, maximum likelihood.

**Table 2** Crown node mean ages obtained with treePL for the main *Carex* lineages

| Lineage | Crown node mean ages (Mya) | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Constrained singletons topology – 10 fossils | Constrained singletons topology – 3 fossils | Singletons query unconstrained tree – 10 fossils | Singletons query unconstrained tree – 3 fossils | Escudero et al. (2012)[a] | Spalink et al. (2016)[b] | Léveillé-Bourret et al. (2018b)[c] | Márquez-Corro et al. (2019)[d] | Uzma et al. (2019)[e] |
| *Carex* | 37.17 | 34.17 | 37.21 | 34.64 | 42.19 | 33.9 | 31.33 | 34.52 | – |
| subg. *Siderosticta* | 24.64 | 14.86 | 24.79 | 16.81 | 5.8 | 10.8 | 22.10 | 15.45 | 17.2 |
| subg. *Psyllophorae* | 24.41 | 21.17 | 24.91 | 17.37 | 10.36 | – | 18.80 | 14.76 | 19.4 |
| subg. *Euthyceras* | 23.06 | 17.41 | 29.10[f] | 19.85[f] | – | – | 18.55 | 16.68 | 24.0 |
| "Caricoid clade" (*Psyllophorae* + *Euthyceras*) | 24.50 | 21.17 | – | – | 23.07 | 14.5 | 23.67 | 21.55 | – |
| subg. *Uncinia* | 22.91 | 15.50 | 22.91[f] | 15.64[f] | 7.85[f] | – | 3.73[f] | – | – |
| subg. *Vignea* | 23.29 | 18.64 | 22.28 | 22.98 | 24.05 | 12.2 | 20.17 | 20.10 | 23.8 |
| subg. *Carex* | 25.18 | 23.21 | 32.01 | 30.16 | 21.14 | 16.1 | 23.12 | 22.37 | 22.9 |

Results from the four dating analytical approaches are shown, whether constraining or not the topology with the Hyb-Seq tree (Villaverde et al., in review), and enforcing 10 vs. three fossils priors (see Table 1) as primary calibrations points. Age estimates from previous studies dating *Carex* are also given for comparison; [a]One *Carex* (*C. tsagajanica*) fossil calibration; [b]Six Cyperaceae (one *Carex*) fossil calibrations; [c]Three *Carex* + seven other Cyperaceae + six other Poales fossil calibrations; [d]Eight Cyperaceae secondary + three *Carex* fossil calibrations; [e]One *Carex* (*C. colwellensis*) fossil calibration; [f]subg. *Uncinia* members nested within subg. *Euthyceras*.

**Table 3  A,** Summary of the number of dispersal events (and standard deviations) between the different considered regions (sources of dispersal in rows and sinks of dispersals in columns) inferred by the BSM analysis under the DEC model. Cell color indicates the range of the number of inferred dispersal events: blue, 0–5; green, 5–10; yellow, 10–30; orange, 30–50; red, >50. All events correspond to anagenetic dispersal (range expansion) events. For the events inferred under the DEC+J model, see Table S1. **B,** Summary of the type of biogeographical events inferred with BSM under the DEC model

**A**

| To \ From | Nearctic | W Palearctic | E Asia | Afrotropic | Neotropic | Pacific | Total (%) |
|---|---|---|---|---|---|---|---|
| Nearctic | – | 44.06 (4.15) | 55.83 (5.17) | 4.43 (1.18) | 46.46 (2.62) | 8.71 (1.71) | 159.49 (37.86) |
| W Palearctic | 26.59 (3.69) | – | 32.43 (4.13) | 10.31 (1.55) | 6.41 (1.84) | 5.39 (1.61) | 81.13 (19.26) |
| E Asia | 49.26 (5.12) | 66.97 (4.62) | – | 8.37 (1.45) | 5.72 (1.92) | 13.93 (1.71) | 144.25 (34.24) |
| Afrotropic | 1.52 (0.67) | 1.36 (0.98) | 1.48 (0.85) | – | 1.04 (0.83) | 0.73 (0.76) | 6.13 (1.46) |
| Neotropic | 6.81 (2) | 2.2 (1.45) | 2.07 (1.39) | 2.14 (0.89) | – | 4.34 (1.3) | 17.56 (4.17) |
| Pacific | 1.58 (1.21) | 1.77 (1.14) | 3.81 (1.42) | 1.93 (0.98) | 3.64 (1.29) | – | 12.73 (3.02) |
| Total (%) | 85.76 (20.36) | 116.36 (27.62) | 95.62 (22.70) | 27.18 (6.45) | 63.27 (15.02) | 33.1 (7.86) | 421.29 (100) |

**B**

| Mode | Type of event | Cladogenetic/Anagenetic event | Mean (SD) | Percentage |
|---|---|---|---|---|
| Speciation within area | Narrow sympatry | Cladogenetic | 1057 (9.21) | 58.33 |
| | Speciation subset | Cladogenetic | 191.3 (10.12) | 10.55 |
| Dispersal | Founder events | Cladogenetic | 0 | 0 |
| | Range expansions | Anagenetic | 421.3 (3.46) | 23.25 |
| Vicariance | Vicariance | Cladogenetic | 142.6 (3.86) | 7.87 |
| Total | | Cladogenetic | 1391 | 76.76 |
| | | Anagenetic | 421 | 23.23 |
| | | | 1812 | 100 |

BSM, biogeographic stochastic mapping; DEC, dispersal-extinction-cladogenesis.

Results from the BSM analyses under the DEC (Table 3) and DEC+J (Table S1) models were similar in terms of the number of inferred biogeographic events (1812 and 1714, respectively), the proportion of events cladogenetic rather than anagenetic (81.15/18.84% and 76.76/23.23%, respectively), and the relative importance of the areas as source/sink of dispersal events. In addition, DEC+J inferred 119.5 founder events, a parameter (J) not included in the DEC model.

Under both models, the great majority of dispersal events were inferred to have taken place between Northern Hemisphere regions (Nearctic, E Asia, and W Palearctic), both as sources (c. 91%) and sinks (c. 71%) of dispersals events (Fig. 2B; Tables 3A, S1A). Southern Hemisphere landmasses play a minor role as the source of dispersals (c. 9%, half of which correspond to dispersal from the Neotropic to the Nearctic) but are more relevant as sinks for dispersals (c. 28%; Fig. 2B; Tables 3A, S1A). In addition, when the type of dispersal is considered in the DEC+J model (anagenetic dispersal or range expansion vs. cladogenetic dispersal or founder event), the Southern Hemisphere gains importance as a sink of cladogenetic dispersal events (39.4%; Table S1A). Prevailing dispersal routes (Fig. 2B) are between any of the Northern Hemisphere areas and, to a much lesser extent, between any of these as a source and a Southern Hemisphere area as a sink (except for the dispersal connection between the Nearctic and the Afrotropic, which

is negligible). An additional significant connection was inferred from the Neotropic to the Nearctic. The relative contribution of dispersal connections among Southern Hemisphere areas is negligible when compared to the rest of the world. The areas that were inferred to be the most important sources of dispersal were the Nearctic and E Asia, with c. 38% and c. 35% (considering both DEC and DEC+J models) of the total number of dispersal events inferred to have taken place from each of these areas (Tables 3, S1A). The most important sink for dispersal was the W Palearctic, at 27% of dispersals inferred to have colonized this area. E Asia and the Nearctic act as sources for dispersals more frequently than as sinks. Thus, more than 70% of the total dispersal events were inferred to have taken place from these two areas (see above), whereas less than half of dispersals were inferred to have colonized these areas (c. 20% and c. 23%, respectively). In contrast, the remaining areas act less frequently as source than sink, which is especially evident for the Neotropic (c. 4% vs. c. 15%) and, to a lesser extent, the W Palearctic (c. 20% vs. c. 27%), the Afrotropic (c. 1% vs. c. 6%), and the Pacific (c. 3% vs. c. 8%). Finally, the majority of biogeographical events inferred by BSM correspond to cladogenetic events (77%, 81% in DEC and DEC+J, respectively), of which speciation within an area appears as the most frequent event (69%, 71%), whereas cladogenetic dispersal (7%, only in DEC+J model) and vicariance (3%, 8%) made relatively minor contributions (Tables 3B, S1D). In summarizing the timing and frequency of biogeographic events, the BSMs indicate some uncertainty in the models, particularly at the basal nodes of the tree where all cladogenetic and anagenetic event types were estimated with equal frequency (Fig. S5). Over the past 25 million years, BSM analyses suggest that while within-area speciation is most common, dispersal and vicariance events collectively accounted for 25–60% of all biogeographical events. The frequency of narrow within-area speciation plummeted while the frequency of anagenetic dispersal rises dramatically (Fig. S5). These anagenetic events largely reflect the exchange of lineages amongst the Nearctic, the W Palearctic, and E Asia (Fig. 2B).

### 3.3 Diversification
Net diversification rates across the 10-fossils tree range from 0.138 to 1.041, a 7.5-fold range, under a prior of 1 shift, and from 0.107 to 1.090, a 10.2-fold range, under a prior of 50 shifts. The estimated number of shifts also differs significantly, from 17.49+/−3.33 (SD) under a shift prior of 1 to 27.68+/−4.50 under a prior of 50 shifts. However, these differences do not manifest in obvious discrepancies in the model-averaged net diversification rates as estimated on the tree, which correlate strongly (for net diversification rate, branchwise Pearson product-moment $r = 0.989$, $P < 1 \times 10^{-16}$), so we will restrict our discussion to the single-shift prior (Fig. 3). Moreover, as there are many possible shifts to discuss, we focus on the 13 shifts to net diversification rates above 0.5 species·My$^{-1}$, which were distributed across the tree (Table 4; Fig. 3A). In the Northern Hemisphere, five shifts comprised mainly Nearctic radiations, and four shifts comprised largely circumboreal groups. The mean age of the crown node of clades involved in most shifts affecting Nearctic radiations fell in a window of about 2 My

flanking the Mio-Pliocene boundary (5.3 Mya), except for the American sect. *Acrocystis*, which had a crown age of around 2.6 Mya, at the end of the Pliocene. For the circumboreal groups, the shift happened 7–5.4 Mya, at the end of the Miocene. Two additional shifts happened in groups from E Asia during the Miocene (8.2 Mya; core *Kobresia*) and Pliocene (4.1 Mya; core sect. *Clandestinae*). In the Southern Hemisphere, there are two other additional shifts: New Zealand clade of sect. *Uncinia* (7.5 Mya, Late Miocene) and sect. *Echinochlaenae* (centered in New Zealand, 4.5 Mya, early Pliocene).

## 4 Discussion

### 4.1 Towards a unified interpretation of *Carex* phylogeny: Robustness of the inferences under different approaches
The phylogenies presented in this study represent by far the most comprehensive sampling to date in the two decades of phylogenetic studies of the megadiverse genus *Carex*, with 100% of sections and almost 66% of accepted species included. However, it should be noted that our species sampling is not uniformly distributed geographically: while there are nearly completely sampled geographic areas, such as the Nearctic and the W Palearctic (>97% species), other areas have a high proportion of nonsampled species, particularly E Asia and the Neotropic (c. 51% and 57%, respectively; Data S1; Govaerts et al., 2019). Given the great richness of species in E Asia (more than 1000 species; Fig. 1B), this region appears to be the critical sampling gap that requires filling in future phylogenetic studies. Nonetheless, our complete sampling of sections likely ensures a good coverage of the phylogenetic diversity of the genus.

The multiple tips query tree with 4470 accessions (Data S3) is the largest *Carex* phylogeny hitherto built to date. Our curation procedure, which resulted in the exclusion of almost 20% of all initially gathered accessions (>1000 concatenated ETS–ITS–*matK* sequences), underlines the need for carefully checking the phylogenetic placement of *Carex* sequences, both newly obtained and downloaded from GenBank, to discard possible contaminations or misidentifications and mislabelings. Our analytical approaches for phylogenetic reconstruction, namely the scaffolding approach (GCG, 2016a) and twofold constraint procedure starting with a phylogenomic (Hyb-Seq) backbone tree (Villaverde et al., in review), proved successful in overcoming some of the problems posed by the large amount of missing data in our dataset (60% and 41% in the multiple tips and singletons matrices, respectively) and yielded relationships largely in agreement with previous studies (e.g., GCG, 2016a) for our greatly expanded species sampling. While three gene regions are certainly insufficient to resolve many of the relationships in the genus, our twice-constrained approach ensures that the deeper relationships are driven by the Hyb-Seq nuclear topology, while the species relationships towards the tips are primarily given by the three barcode DNA regions (ETS, ITS, and *matK*), where they provide the best resolution. While the large amount of missing data in our matrices may results in errors in some of the shallowest species-level phylogenetic relationships, remarkably we were able to retrieve a mostly highly supported topology based on only three DNA regions
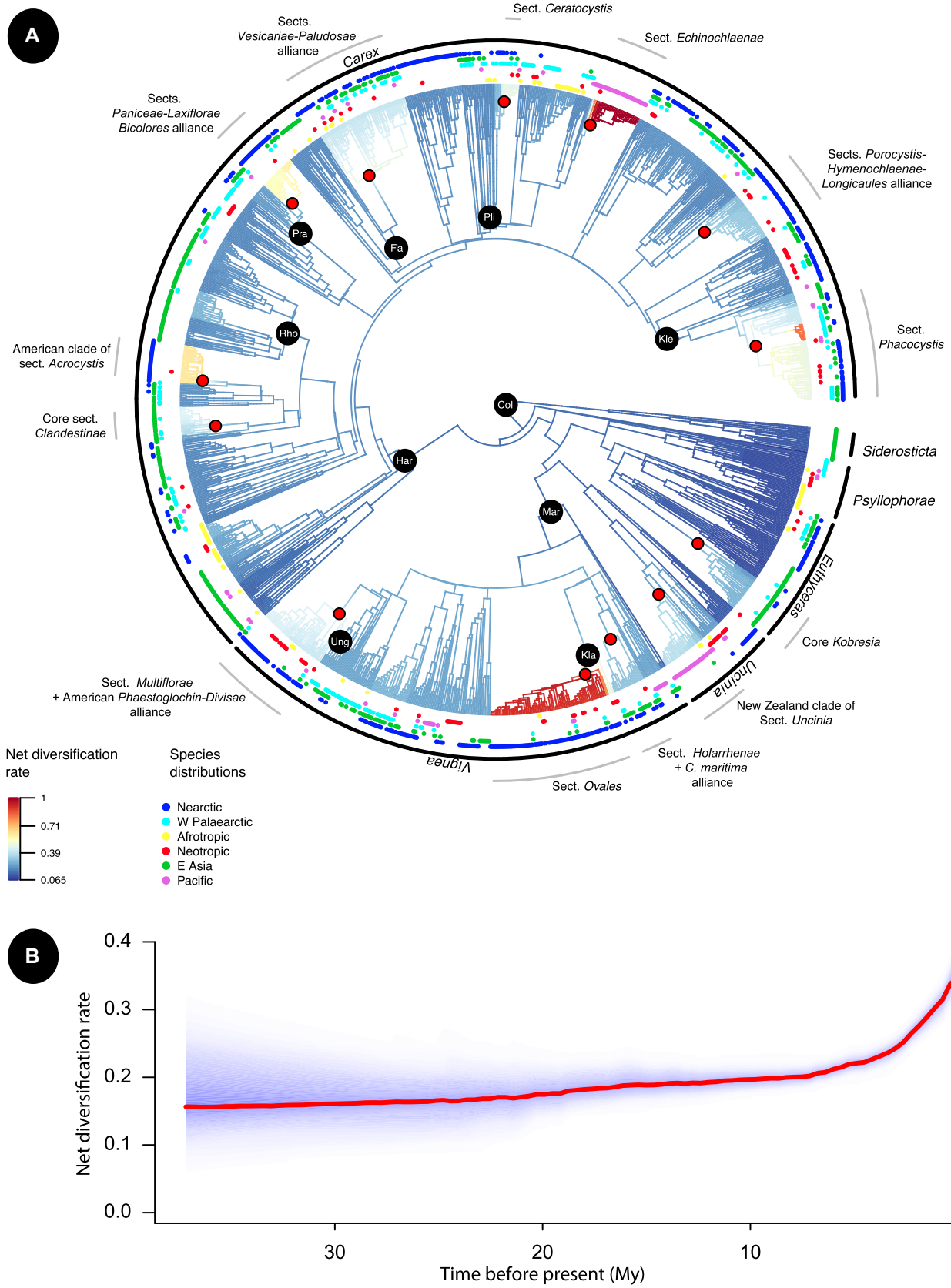
www.jse.ac.cn

*J. Syst. Evol.* 57(6): 695–718, 2019

**Fig. 3.** Continued

**Table 4** Clades involved in the diversification shifts detected by BAMM analysis, with their distribution, and mean age of crown node according to the constrained singletons tree calibrated with the 10 available fossils (Data S6)

| Clade | Distribution | Mean age of crown node (Mya) |
|---|---|---|
| **Subg. *Euthyceras*** | | |
| Core *Kobresia* | E Asia (Himalayas) | 8.2 |
| **Subg. *Uncinia*** | | |
| New Zealand clade of sect. *Uncinia* | Primarily New Zealand | 7.5 |
| **Subg. *Vignea*** | | |
| Sect. *Holarrhenae* + *C. maritima* alliance (excl. sect. *Chordorrhizae*) | Primarily circumboreal | 6.1 |
| Sect. *Ovales* (incl. *C. bonplandii* clade) | Almost entirely North American | 5.4 |
| Sect. *Multiflorae* + American sects. *Phaestoglochin-Divisae* alliance | Primarily North American | 6.1 |
| **Subg. *Carex*** | | |
| Core sect. *Clandestinae* | E Asia (Himalayas) | 4.1 |
| American clade of sect. *Acrocystis* (excl. *C. pilulifera*) | Almost entirely North American | 2.6 |
| Sects. *Paniceae-Laxiflorae-Bicolores* alliance (excl. *C. olbiensis*) | Almost entirely North American | 4 |
| Sects. *Vesicariae-Paludosae* alliance (excl. *C. aureolensis* clade) | Primarily circumboreal | 7 |
| Sect. *Ceratocystis* | Primarily circumboreal | 5.4 |
| Sect. *Echinochlaenae* (excl. *C. blakei* clade) | Almost entirely New Zealand | 4.5 |
| Sect. *Porocystis-Hymenochlaenae-Longicaules* alliance (excl. *C. sartwelliana* clade) | Primarily North American | 5.8 |
| Sect. *Phacocystis* (sister to *C. podocarpa* clade) | Circumboreal, with a nested diversification rate increase in a primarily W Palearctic clade | 6 (nested shift around 0.9) |

Clades are presented in clock-wise order from the tree root according to Fig. 3A; BAMM, Bayesian analysis of macroevolutionary mixtures.

(ETS, ITS, and *matK*), with clade supports above 80% BS and 90% SH for most internal nodes in both, the singleton trees (Data S4 and S5) and the multiple tips query tree (Data S3). That said, a broader sampling of loci will be necessary to recover a better-supported species tree that accounts for among-gene incongruence that we cannot resolve with only the aforementioned nrDNA and cpDNA loci.

Our dating strategy took advantage of the recent assessment of the rich *Carex* fossil record (Table 1; Jiménez-Mejías et al., 2016b) and compared various analytical approaches to explore the sensitivity of our age estimates to alternative topologies and calibration schemes. Estimated ages using 10 fossils as calibration points were considerably older than ages based on only three fossils (Table 2). This agrees with other studies exploring the effect of the number, quality and age of calibration points on divergence estimates

(Sauquet et al., 2012; Tripp & McDade, 2014; Saladin et al., 2017) and advocates for the use of as many reliable fossil calibrations as possible (Hug & Roger, 2007), especially on deeper nodes (Mello & Schrago, 2014). Estimated divergence times are generally older than those obtained in previous studies dating *Carex* (Escudero et al., 2012; Spalink et al., 2016b; Léveillé-Bourret et al., 2018b; Márquez-Corro et al., 2019; Uzma et al., 2019), which relied on a limited sampling of the genus and/or used few reliable primary calibration points (Table 2). As an exception, the use of a controversial *Carex* fossil from the Early Paleocene (*C. tsagajanica* Krassilov) to constrain the stem node of *Carex* yielded a slightly older *Carex* crown age (mean 42.19 Mya) in Escudero et al. (2012) relative to our estimates. However, significant questions regarding the identity of this fossil have been raised by Jiménez-Mejías et al. (2016b), so we discarded it and used *C.*

◀

**Fig. 3. A,** Phylorate plot obtained from the analysis of diversification rate in *Carex* with BAMM and based on the ML constrained singletons chronogram calibrated with 10 fossils (Data S6). Tree branch color indicates the model-averaged net diversification rates (species per Million year; spp/My) along the branches. Thirteen diversification rate shifts to rates above 0.5 spp/My are marked in the figure using red circles. Black circles indicate fossil calibrations used in the dating analysis (Table 1): *C. colwellensis* (Col), *C. marchica* (Mar), *C. ungeri* (Ung), *C. klarae* (Kla), *C. hartauensis* (Har), *C. praehirta* (Pra), *C. klettvicensis* (Kle), *C.* sect. *Rhomboidales* (Rho), *C. plicata* (Pli), and *C. flagellata* (Fla). Encircling text and lines depict (from inside to outside): geographic distribution of species (colored dots according to legend and Fig. 2B), subgenera (black arcs), and clades in which diversification rate shifts were detected (according to Table 4). **B,** Net diversification rate through time plot for *Carex*. The thick red line marks mean net diversification rate, and the shaded range indicates its rjMCMC confidence interval. BAMM, Bayesian analysis of macroevolutionary mixtures; ML, maximum likelihood; rjMCMC, reversible jump Markov chain Monte Carlo.

*colwellensis* (Table 1) from the Late Eocene as the oldest reliable *Carex* fossil available to constrain the crown node of *Carex*.

## 4.2 The early biogeographic history of *Carex*

All biogeographic reconstructions clearly inferred E Asia as the unambiguous ancestral area not only for the whole genus but also for most of its main lineages. Exceptions include subg. *Psyllophorae* and subg. *Uncinia*, which arose in the W Palearctic and America respectively (Fig. 2A). The E Asian origin of *Carex* has been proposed since the discovery that subg. *Siderosticta*, formed exclusively of E Asian species, is sister to the rest of the genus (Waterway et al., 2009; Starr et al., 2015). The subsequent discovery of tribe Sumatro-scirpeae, another SE Asian lineage, as sister to *Carex* (Léveillé-Bourret et al., 2018a, 2018c; Semmouri et al., 2019), reinforced this hypothesis. However, the early biogeographic history of *Carex* has only been tested with a limited sampling and within a broader study focused on all Cyperaceae (Spalink et al., 2016b). Our study lends extensive support to the general "out-of-Asia" biogeographic pattern previously suggested for *Carex*. Given that our main sampling gaps in *Carex* also correspond to E Asian species (see above), an increased sampling in E Asia will probably reinforce this area as both the diversification origin and main species diversity center of the genus. Additional sampling in E Asia may of course also reveal additional reversions to Asia from other regions. Thus, E Asia could be considered not only the "cradle" of the early diversified lineages, but also a "museum" (Moreau & Bell, 2013) of *Carex* species, as inferred in other important plant groups with high diversity of extant species (Dupin et al., 2016; Echeverría-Londoño et al., 2018; Huang et al., 2019).

However, our results should still be interpreted carefully given that ancestral area estimation can be strongly biased by variation in extinction rates between different geographic areas. When this occurs, areas with the lowest extinction rate are often mistakenly inferred as ancestral even when they are not (Sanmartín & Meseguer, 2016). Eastern Asia is the region of the Northern Hemisphere that has been least impacted by both Pleistocene glaciations and previous Cenozoic climatic changes (Milne & Abbott, 2002; Manchester et al., 2009). It is thus likely that extinction rates have been lower in Eastern Asia compared to other Northern hemisphere areas over most of the evolutionary history of *Carex*, and this could bias ancestral area estimation in a way that cannot be solved with additional taxonomic sampling. Future studies will need to examine the potential role that unequal extinction rates could have played in the diversification and geographic partitioning of *Carex* diversity across the Northern Hemisphere.

Subsequent to its split from its sister group *Sumatro-scirpus*, the earliest lineages of *Carex* persisted exclusively in E Asia for more than 10 million years, before diversification in the Late Eocene (*Carex* stem and crown nodes: 37.2–39.8 Mya; Fig. 2A; Table 2) and the mostly simultaneous diversification of the main lineages in the Late Oligocene (crown nodes: 22.9–25.2 Mya; Table 2). However, the use of the *C. colwellensis* fossil constraint seems to pose a conflict with respect to our dated biogeographic inferences, since it

is reported from the late Eocene (38.0–33.9 Mya) in England (Chandler, 1963). Its age would therefore predate any of our inferred colonization events to the W Palearctic which did not take place at least until the origin of subg. *Psyllophorae* in the late Oligocene (24.4 Mya). This could either indicate migration out of E Asia of ancient *Carex* lineages that are now extinct, or that E Asian endemic lineages were formerly more widespread in Eurasia, and have now become extinct in the W Palearctic but survived in E Asia. Such an inference would mirror the history of other Cyperaceae groups present in the fossil record of Europe but now entirely absent from the continent (e.g., Mapanioideae, Smith et al., 2009; *Dulichium* Pers., Mai & Walther, 1988), and also argues for a careful interpretation of our inferred biogeographic history. Another hypothesis that must be considered with caution is that the age of *Carex* may be much older than *C. colwellensis*, and in that case, we would be underestimating the age of the genus. This idea is plausible given that Asia, the inferred ancestral area of *Carex*, has fewer fossil records in general because of the smaller number of palaeobiological studies focused in this area (see Jiménez-Mejías et al., 2016b).

One of the most interesting discoveries of this work is the near-simultaneous diversification detected in the main *Carex* lineages (i.e., subg. *Carex*, *Euthyceras*, *Siderosticta*, and *Vignea*; crown nodes: 23.1–25.2 Mya; Table 2) in E Asia during the late Oligocene (Chattian; 28.1–23.03 Mya) and their subsequent expansion (Figs. S1, S2), which constitute a remarkable example of biogeographic congruence for extant species groups (shared geographic and temporal scenario; Vargas et al., 2014). While this near-simultaneous diversification of the clades that gave rise to the world's *Carex* diversity has not previously been noted, this major diversification of *Carex* has been attributed both to global cooling from the late Eocene to the Miocene, as well as a shift in the mode of chromosome evolution (Escudero et al., 2012; Márquez-Corro et al., 2019). Such combinations of abiotic and biotic drivers are increasingly found implicated in key angiosperm diversifications (Bouchenak-Khelladi et al., 2015; Fernández-Mazuecos et al., 2018; Otero et al., 2019).

Two contrasting patterns are detected in the largely Southern Hemisphere radiations of subg. *Psyllophorae* and subg. *Uncinia*. On one hand, the South American origin of sect. *Uncinia* (Figs. S1, S2) and its increased diversification rate (Fig. 3A; Table 4) suggest a joint effect of the colonization of the New Zealand via long-distance dispersal (LDD; Uribe-Convers & Tank, 2015) together with a potential key innovation, the presence of a hooked rachilla, one of the only two unequivocal epizoochoric syndrome known in the genus (see Villaverde et al., 2017a; the other being the hooked utricle beaks of *C. collinsii*, Reznicek, pers. obs.). On the other hand, subgenus *Psyllophorae* has been found to be of W Palearctic origin (Fig. 2A) with a subsequent Early Miocene (17.2 Mya) dispersal to the Southern Hemisphere (Data S6; Figs. S1, S2). The long branches and deep nodes of subg. *Psyllophorae* suggest either differential extinction or low diversification rates in that group (Fig. 2A; Data S6). Species of that clade lack any clear epizoochorous trait (except perhaps the short-protruding rachilla of *C. camptoglochin*). Nonetheless, this clade shows a remarkable distribution that includes disjunctions between and within Northern and Southern Hemispheres that are congruent with both Rand

Flora (Pokorny et al., 2015; Mairal et al., 2017) and Gondwanan patterns (Givnish & Renner, 2004; Sanmartín & Ronquist, 2004).

Congruent with the different number and diversity of species that have been reported from the different landmasses (Fig. 1B), BSM revealed a strong asymmetry in the number of inferred dispersal events between the Northern and Southern Hemispheres (Fig. 2B). The three Northern Hemisphere landmasses (E Asia, Nearctic and, to a lesser extent, W Palearctic) clearly emerge as the most important sources and destinations of colonization (Fig. 2B; Tables 3A, S1A). On the other hand, primarily Southern Hemisphere regions (Neotropic, Afrotropic, and Pacific region) are retrieved as colonization sinks (although with a much-reduced numbers of colonizations inferred than for Northern Hemisphere regions), while their role as colonization sources is almost negligible (exceptions including a few colonizations inferred from the Neotropic to the Nearctic; Fig. 2B; Tables 3A, S1A). Asymmetric patterns of dispersal between different continents have been previously reported from other groups (e.g., Sanmartín et al., 2007; Dupin et al., 2016; Zuloaga et al., 2018). While animal and plant migration from the Late Miocene (c. 7 Mya) onwards is considered about 30% more frequent from South America to North America (Bacon et al., 2015), our results inferred 5–7 times as many dispersals in the opposite direction (Fig. 2B; Tables 3A, S1A), which is congruent with the general pattern of the amphitropical American disjunctions (Simpson et al., 2017). The positive relationship between species richness in each region (Govaerts et al., 2019; Fig. 1B) and the number of inferred dispersal events (Fig. 2B; Tables 3A, S1A) should be taken into account as it regards this asymmetric dispersal pattern (Dupin et al., 2016). In any case, the well-known faunistic turnover in South America due to the Great American Biotic Interchange (Stehli & Webb, 1985) contrasts with the relative stability of the Neotropical flora, which even migrated northwards into Central and North America as climatic conditions changed and migrations routes opened (Willis et al., 2014; Willis & Davis, 2015). The colonization of boreo-temperate elements into the Tropics, such as the one shown by *Carex*, has only been reported in few cases (e.g., in the Afrotropic, Escudero et al., 2009; Gehrke & Linder, 2009; Míguez et al., 2017; or the Neotropic, Uribe-Convers & Tank, 2015; Simpson et al., 2017).

The inferred timing of biogeographic events (Fig. S5) strongly suggests that multiple, recurrent LDD events better explain *Carex* biogeographic patterns than tectonic vicariance or Northern Hemisphere land-bridge hypotheses, perhaps with the exception of some relatively recent Beringian lineages (see Maguilla et al., 2018). The occurrence of dispersal events is clearly biased towards more recent times, especially to the last 10 My corresponding to the most recent 25% of the age of the genus. LDD has often been invoked to explain widely disjunct ranges in *Carex* and Cyperaceae as a whole (e.g., Escudero et al., 2009; Viljoen et al., 2013; Gebauer et al., 2015; Spalink et al., 2016b; Míguez et al., 2017; Villaverde et al., 2017a). A skewed distribution of dispersal events towards present times has also been observed in other groups in which LDD is regarded as a critical process shaping their biogeographic patterns (Tripp & McDade, 2014; Dupin et al., 2016; Ruhfel et al., 2016; Rose et al., 2018; Huang et al., 2019). With

respect to the dispersal source, recurrent dispersal from E Asia is inferred not to have started until the late Oligocene-early Miocene (c. 25–20 Mya; Figs. 2A, S1, S2; Data S6), consistent with a temporal lag between the origin of the group and the diversification of its main lineages. This has been also found in other groups (Dupin et al., 2016; Zuloaga et al., 2018; Huang et al., 2019), although extinct lineages may have dispersed earlier from E Asia as noted above.

In situ cladogenetic diversification is inferred as the most important biogeographic event, accounting for about 70% of the total events inferred under both the DEC and DEC+J models (Tables 3B, S1D). This is explained by the large size of the areas coded for the biogeographical analyses (Fig. 2B) and the presence of large clades endemic to an area (Figs. 2A, S1, S2). As in other studies (Berger et al., 2016; Dupin et al., 2016; Spalink et al., 2016a, 2016b; Zuloaga et al., 2018), the real importance of allopatric speciation can be masked by these facts, and thus caution should be taken in interpreting our results, that suggest within-area speciation, but not sympatric speciation in the strict sense, as predominant in *Carex*. Biogeographic analyses focusing on particular clades and implementing a fine-scale geographic division within the wide areas considered here would probably unveil a greater role of allopatric speciation processes (i.e., founder events; Spalink et al., 2016a; Johnson et al., 2017).

### 4.3 Biogeographic diversity at shallow evolutionary scales
As mentioned previously, the enormous diversity in *Carex* and its worldwide distribution make it possible for the group to display almost every distribution pattern known in angiosperms (Raymond, 1955). This fact, combined with its ecological importance and ubiquity, make biogeographic insights in *Carex* relevant beyond the limits of the genus. In this epigraph, we comment on some of the most striking and less-explored aspects of *Carex* biogeography and discuss possible drivers of the observed geographical ranges. Remarkably, all of the biogeographic events discussed here transpired in the late Miocene–Pliocene, a geological epoch of marked global cooling (Gradstein et al., 2004) that probably contributed to the geographical expansion of the involved lineages.

While many shallow clades (sectional level and below) are restricted to a single area, a number are widespread (Figs. S1, S2), with members native to up to five or even all six coded areas. It is noteworthy that these widespread groups mostly belong to subg. *Carex* (e.g., *C. pseudocyperus* and allies, sects. *Ceratocystis*, *Spirostachyae-Echinochlaenae* alliance and *Phacocystis*) and, to a lesser degree subg. *Vignea* (sect. *Glareosae*) and subg. *Euthyceras* (sect. *Capituligerae*). While this may just be ascertainment bias—the probability of detecting any significant clade-level effects increases in larger clades, and subg. *Carex* and *Vignea* are by far the largest—it may also suggest a dispersal and colonization ability unique to these groups, as discussed in previous studies (Escudero et al., 2009; Jiménez-Mejías et al., 2012b; Villaverde et al., 2015b, 2017b; Maguilla et al., 2018). However, none of these groups identified as geographically widespread display clear anemochorous or epizoochorous traits (though the bent utricle beaks in some sect. *Ceratocystis* species and the spreading utricle beak teeth in *C. pseudocyperus* and allies may play a role in animal dispersal). Such large distributions

are particularly remarkable in light of the putatively non-specialized diaspore in most *Carex* groups. Different hypotheses have been proposed to explain the dispersal ability of *Carex*. One of the most widely accepted is the possible endozoochorous dispersal of the seeds by birds, but epizoochory of diaspores carried on birds' feet cannot be definitely ruled out (see discussion in Villaverde et al., 2017a). In both cases, the two kinds of animal dispersal would be facilitated by the small size of most *Carex* fruits and the wetland habitat of a majority of the species (including the six widespread groups mentioned above).

While the predominant dispersal routes between the landmasses along the Atlantic mainly involve the well-known (North) amphi-Atlantic pattern (Nearctic-W Palearctic; Hultén, 1958) and North-South relationships (Nearctic–Neotropic, and W Palearctic-Afrotropic; Fig. 2B), there is a striking crossed pattern in certain groups of species, involving dispersals between the W Palearctic and the Neotropic (e.g., sects. *Abditispicae-Thuringiacae* alliance, *C. punctata* and *C. extensa* groups in sect. *Spirostachyae*; Figs. S1, S2; Table 3A), and, almost incidentally, between the Nearctic and the Afrotropic (*C. conferta* group, perhaps also *C. pseudocyperus*; Figs. S1, S2). Phylogenetic reconstructions (i.e., Data S5) strongly support the nested position of these Neotropical and Afrotropical groups within predominantly W Palearctic and Nearctic lineages, respectively. Of these, the oldest disjunction is the one involving sects. *Abditispicae-Pellucidae*, dating back to the late Miocene (8.4 Mya), while the others are entirely placed in the Pleistocene (1.9–1.7 Mya) (Figs. S1, S2). Escudero et al. (2009) previously discussed the W Palearctic origin of sect. *Spirostachyae* and subsequent colonization of SE South America and the Southern Atlantic archipelago of Tristan da Cunha involving LDD. Without any known bird flyway among the involved landmasses, it seems to point to nonstandard stochastic dispersal processes (Nogales et al., 2012) as the origin of these LDDs between nonadjacent landmasses. In this sense, palaeo-paths of the trade winds may help explaining this striking crossed pattern (see McGee et al., 2018).

While the vast majority of *Carex* species have a preference for cold-temperate climates, shifting to high elevations in the Tropics, a few species-rich groups dwell in tropical montane forests (e.g., sect. *Fecundae*, most species of subg. *Siderosticta*, or the Tropical African clade of sect. *Spirostachyae*). Of them, the Asian-African-American sect. *Indicae* (hereafter the "AAA clade"), is the only one that exhibits a pantropical distribution. In addition, the AAA clade is the most successful in terms of species diversity (c. 80) and dispersal/colonization ability. Originating in tropical E Asia, the species comprising the AAA clade dispersed to the Afrotropics at the end of the Miocene (5.64 Mya) and from there dispersed across the Atlantic to tropical America, apparently in a single Pliocene (2.65 Mya) LDD (Figs. S1, S2; Data S6).

### 4.4 Diversification rate shifts in the light of biogeography
The starkly different pattern of species richness and extent of distribution between the near cosmopolitan *Carex* with 2000 species and its sister group *Sumatroscirpus* with three species that are restricted to SE Asia (Léveillé-Bourret et al., 2018a) is congruent with a positive area-richness correlation pattern and suggests that colonization out of E Asia may have

facilitated the diversification in *Carex* through the occupation of novel ecological niches (Spalink et al., 2016b). A similar scenario is found within *Carex*, with the subg. *Siderosticta* being comprised by only c. 28 E Asian species, while its sister group, the rest of the genus, comprises almost 99% of all *Carex* species (>1950 spp) and is subcosmopolitan. In fact, when compared to the rest of Cyperaceae, a diversification rate shift in *Carex* has been previously detected involving all the non-Siderosticta clade (Escudero et al., 2012; Spalink et al., 2016b). This transition also entails a shift in the mode of chromosome evolution from low rates of chromosome evolution (including polyploidy -genome duplication- and dysploidy -fission and fusion of chromosomes-) to drastically increased rates of dysploidy (Hipp et al., 2009; Escudero et al., 2012; Márquez-Corro et al., 2019). As a result, teasing apart the relative effect of chromosome evolution and ecological diversification on sedge diversity may prove difficult. That said, our analyses suggest that the success of *Carex* globally may have depended on colonization of novel geographic areas, where over and over alternative axes of niche space were available for diversification (Wellborn & Langerhans, 2015).

The particular groups where diversification shifts have been detected (Table 4) seem to share certain features of sympatry and synchrony (see below) correlated with geoclimatic cooling phenomena during the Late Neogene (Late Miocene-Pliocene), as previously found at smaller taxonomic and geographical scales (Gebauer et al., 2014; Hoffman & Gebauer, 2016; Hoffmann et al., 2017). This could have been favored by the intrinsic cold-adapted nature of *Carex*, as suggested previously by Escudero et al. (2012). In the Northern Hemisphere, for example, diversification shifts involved primarily two kinds of radiations: predominantly Nearctic, and predominantly circumboreal (five and four radiations respectively; Table 4; Fig. 3A). The diversification boost promoted by global cooling at the Late Miocene-Pliocene likely had greater impact on Northern Hemisphere clades simply because of the historically higher northern *Carex* diversity compared to the Southern Hemisphere.

The contrasting diversification pattern on both sides of the Atlantic, with five shifts (Table 4) located in the Nearctic, but only one centered exclusively in the W Palearctic (core sect. *Phacocystis*), demands further explanations involving historical climatic differences between the North Atlantic landmasses. These contrasting climatic patterns could date back to the closure of the isthmus of Panama, which initiated with the onset of the Pleistocene glaciations (e.g., Bartoli et al., 2005). The conjoined effect of this climatic transformation, together with the mass Pleistocene extinctions in Europe (favored by the E-W orientation of mountain ranges and the Mediterranean sea, and by disappearance of temperate forest; Svenning, 2003) could help to explain not only the contrasting diversification patterns but also the remarkably different diversity levels (c. 562 species in the Nearctic vs. c. 244 in the W Palearctic; Fig. 1B) between landmasses on opposite sides of the North Atlantic.

Also striking are the few diversification shifts detected in E Asia when compared to the Nearctic, despite its almost twofold species diversity (c. 1000 species; Fig. 1B). While this may reflect the sampling bias in our study, the role of this area as a cradle for the genus, characterized by the early diversification and long-term persistence of the main *Carex*

*J. Syst. Evol.* 57(6): 695–718, 2019

www.jse.ac.cn

lineages (see above), may also be instrumental in this pattern. Under this scenario, the accumulation of lineages in E Asia would have progressively filled available niche spaces, preventing by competitive exclusion diversification by emerging groups (Abrams, 1983). The only two E Asian groups that experienced diversification shifts are centered in the Himalayas (core *Kobresia* and core sect. *Clandestinae*; see Dai et al., 2010), each at a different geological time (Table 4). The former genus *Kobresia* as a whole is believed to have entered the Himalayas about 20.6 Mya (Uzma et al., 2019), which agrees with the mean crown ages retrieved by our dating analyses (21.5 Mya; Data S6). That colonization predates by about 12 Mya the detected diversification shift at the Late Miocene (8.2 Mya), a period where a major uplift of the Himalaya and neighboring mountains has been reported (Zhisheng et al., 2001) together with cooling on the Qinghai-Tibetan Plateau (Favre et al., 2015). This uplift could have created new ecological opportunities enabling the diversification of the already existing core *Kobresia* lineages. Indeed, at a finer geographical scale, core *Kobresia* species could have prevented the diversification of most of the numerous groups that inhabit the Himalayas by the aforementioned competitive exclusion (Uzma et al., 2019). Only core sect. *Clandestinae* species would have diversified in the Himalayas at the Plio-Pleistocene boundary, perhaps promoted by Pleistocene glaciations. In any case, given the sampling gap of E Asian species, additional diversification shifts may be detected in other Asian groups.

In the Southern Hemisphere, there are two additional shifts (Table 4) in two distantly related lineages: New Zealand clade of sect. *Uncinia* at the Late Miocene (7.5 Mya), and sect. *Echinochlaenae*, starting in the Pliocene (4.5 Mya) and entirely centered in New Zealand (Figs. 3A, S1, S2; Data S6). Diversification of sect. *Uncinia* in South America can be linked with the widening of the Drake and Tasmanian Passages, reinforcement of the circumantarctic current, and the freezing of Antarctica (Cantril & Poole, 2012). There is a qualitative change around the Middle Miocene with the consolidation of the Antarctic ice sheet and disappearance of the last tundra remains there (Lewis et al., 2008), coupled with climate cooling and changes in biota composition in the surrounding lands (Mildenhall, 1980; Iglesias et al., 2011; Pole, 2014) that might have facilitated the establishment and diversification of new cold-temperate lineages. Later, the colonization and radiation of both groups within the SE Pacific has been dated by our biogeographic reconstructions as synchronous, happening during the Late Miocene (crown nodes c. 7.5 Mya; Data S6). In the case of sect. *Uncinia*, dispersal was likely from South America whereas in sect. *Echinochlaenae*, remarkably, it has been recovered to have happened from the W Palearctic (Figs. S1, S2), although the possibility of an extinct or unsampled ancestor present in an adjacent area cannot be completely excluded. These lineages could represent two remarkable evolutionary radiations (Simões et al., 2016) in New Zealand, with sect. *Echinochlaenae* (with 40 endemic species) and sect. *Uncinia* (34 endemic species; Ford, 2007; Schönberger et al., 2017) accounting for about 65% of native *Carex* species there. *Carex* is the second most species-rich angiosperm genus in New Zealand, where it also displays an extraordinary endemicity rate (c. 86%; Schönberger et al., 2017). This is

especially interesting from an evolutionary and biogeographical point of view, taking into account its location in the relatively *Carex*-poor Southern Hemisphere, its insular condition and its prolonged geographic isolation for more than 50 million years (Veevers et al., 1991; Schellart et al., 2006).

### 4.5 Final remarks: *Carex* as a case study in how to colonize (almost) the entire planet

Our study contributes significantly to a better understanding of the worldwide macroevolutionary success of *Carex*, which should inform subsequent research on the evolution, biogeography, and ecology of this megadiverse genus. If there are a discrete suite of take-away messages from our study, we consider them to be:

1. **East Asia is the cradle of the genus: *Carex* originated there, and its major clades diversified there synchronously.** Early diversification of subgeneric lineages took place from the Late Eocene to the Late Oligocene, although fossil evidence seems to point to an earlier expansion of now extinct lineages. Synchronous diversification of most main lineages in E Asia constitutes a remarkable example of biogeographic congruence that needs further research.

2. **Modern distribution of *Carex* is primarily a product of recent diversification on northern landmasses.** While crown diversification transpired early in the evolution of the genus, recent speciation events produced most of the species we observe today: diversification rate shifts observed in our study (Fig. 3) affect mostly shallow lineages and are clustered around the late Miocene-Pliocene, possibly promoted by different large-scale climate cooling events that happened during the late Neogene. East Asia, the Nearctic and, to a lesser extent, the W Palearctic, predominate both as dispersal sources and sinks, while landmasses of the Southern Hemisphere act mostly as the destination of colonizations.

3. **Ecological opportunity may have played a large role in *Carex* diversification.** Asymmetric patterns of diversification among regions suggest that the availability of novel niche space after successful colonization of new territories may facilitate *Carex* diversification, as happened in the Nearctic; on the contrary, the accumulation of lineages filling the available niche spaces could have prevented emerging groups from undergoing diversification by competitive exclusion, as inferred in E Asia.

4. **Long-distance dispersals have played a large role in *Carex* distribution.** This is particularly surprising in light of the lack of obvious LDD syndromes in the genus, except for sect. *Uncinia* and *C. collinsii*.

We have yet to understand the elusive question of why *Carex* diversity is apportioned as it is across the globe. Our work, however, provides a foundation needed for understanding patterns of biogeography and diversification in the genus, and lays a groundwork for future studies aimed at understanding what has shaped the disparate patterns of diversity we observe in *Carex*.

## Acknowledgements

## References

Abrams P. 1983. The theory of limiting similarity. *Annual Review of Ecology, Evolution, and Systematics* 14: 359–376.

Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology* 55: 539–552.

Anisimova M, Gil M, Dufayard JF, Dessimoz C, Gascuel O. 2011. Survey of branch support methods demonstrates accuracy power, and robustness of fast likelihood-based approximation schemes. *Systematic Biology* 60: 685–699.

Bacon CD, Silvestro D, Jaramillo C, Smith BT, Chakrabarty P, Antonelli A. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences USA* 112: 6110–6115.

Ball PW. 1990. Some aspects of the phytogeography of *Carex*. *Canadian Journal of Botany* 68: 1462–1472.

Ball PW, Reznicek AA. 2002. *Carex* [Generic description and key to species]. In: Flora of North America Editorial Committee eds. *Flora of North America*. New York, NY and Oxford, UK: Oxford University Press. 23: 254–273.

Bartoli G, Sarnthein M, Weinelt M, Erlenkeuser H, Garbe-Schönberg D, Lea DW. 2005. Final closure of Panama and the onset of Northern Hemisphere glaciation. *Earth and Planetary Science Letters* 237: 33–44.

Benítez-Benítez C, Escudero M, Rodríguez-Sánchez F, Martín-Bravo S, Jiménez-Mejías P. 2018. Pliocene-Pleistocene ecological niche evolution shapes the phylogeography of a Mediterranean plant group. *Molecular Ecology* 27: 1696–1713.

Benítez-Benítez C, Míguez M, Jiménez-Mejías P, Martín-Bravo S. 2017. Molecular and morphological data resurrect the long neglected *Carex laxula* (Cyperaceae) and expand its range in the western Mediterranean. *Anales del Jardín Botánico de Madrid* 74: e057.

Berger BA, Kriebel R, Spalink D, Sytsma KJ. 2016. Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. *Molecular Phylogenetics and Evolution* 95: 116–136.

Berger SA, Krompass D, Stamatakis A. 2011. Performance, accuracy, and web server for volutionary placement of short sequence reads under maximum likelihood. *Systematic Biology* 60: 291–302.

Bouchenak-Khelladi Y, Onstein RE, Xing Y, Schwery O, Linder HP. 2015. On the complexity of triggering evolutionary radiations. *New Phytologist* 207: 313–326.

Brummitt RK. 2001. *World geographical scheme for recording plant distributions*. 2nd ed. Pittsburgh: Hunt Institute for Botanical Documentation.

Cantril D, Poole I. 2012. *The vegetation of Antarctica through geological time*. Cambridge: Cambridge University Press.

Chandler MEJ. 1963. Revision of the Oligocene floras of the Isle of Wight. *Bulletin of the British Museum (Natural History). Geology* 6: 321–384.

Chater AO. 1980. *Carex* L. In: Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA eds. *Flora Europaea, Alismataceae to Orchidaceae*. Cambridge: Cambridge University Press. 5: 290–323.

Dai LK, Liang SY, Zhang SR, Tang YC, Koyama T, Tucker, GC. 2010. *Carex* L. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 23: 285–461.

Derieg NJ, Sanguamphai A, Bruederle LP. 2008. Genetic diversity and endemism in North American *Carex* section *Ceratocystis* (Cyperaceae). *American Journal of Botany* 95: 1287–1296.

Dragon J, Barrington DS. 2009. Systematics of the *Carex aquatilis* and *C. lenticularis* lineages: Geographically and ecologically divergent sister clades of *Carex* section *Phacocystis* (Cyperaceae). *American Journal of Botany* 96: 1896–1906.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.

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

Dupin J, Matzke NJ, Sarkinen T, Knapp S, Olmstead R, Bohs L, Smith S. 2016. Bayesian estimation of the global biogeographic history of the Solanaceae. *Journal of Biogeography* 44: 887–899.

Echeverría-Londoño S, Särkinen T, Fenton IS, Knapp S, Purvis A 2018. Dynamism and context dependency in the diversification of the megadiverse plant genus *Solanum* L. (Solanaceae). Available from https://doi.org/10.1101/348961

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

Egorova TV. 1999. *The sedges (Carex L.) of Russia and adjacent states*. St Louis: Missouri Botanical Garden Press.

Elliott TL, Waterway MJ, Davies TJ. 2016. Contrasting lineage-specific patterns conceal community phylogenetic structure in larger clades. *Journal of Vegetation Science* 27: 69–79.

Escudero M, Hipp A, Waterway MJ, Valente L. 2012. Diversification rates and chromosome evolution in the most diverse angiosperm genus of the temperate zone (*Carex*, Cyperaceae). *Molecular Phylogenetics and Evolution* 63: 650–655.

Escudero M, Luceño M. 2009. Systematics and evolution of *Carex* sects. *Spirostachyae* and *Elatae* (Cyperaceae). *Plant Systematics and Evolution* 279: 163–189.

Escudero M, Valcárcel V, Vargas P, Luceño M. 2009. Ecological vicariance and long distance dispersal significance in the diversification of *Carex* sect. *Spirostachyae* (Cyperaceae). *American Journal of Botany* 96: 2100–2114.

Escudero M, Vargas P, Arens P, Ouborg NJ, Luceño M. 2010. The east-west-north colonization history of the Mediterranean and Europe by the coastal plant *Carex extensa* (Cyperaceae). *Molecular Ecology* 19: 352–370.

*J. Syst. Evol.* 57(6): 695–718, 2019

www.jse.ac.cn

Favre A, Päckert, M, Pauls SU, Jähnig SC, Uhl D, Michalak I, Muellner-Riehl AN. 2015. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biological Reviews of the Cambridge Philosophical Society* 90: 236–253.

Fernández-Mazuecos M, Blanco-Pastor JL, Juan A, Carnicero P, Forrest A, Alarcon M, Vargas P, Glover BJ. 2018. Macroevolutionary dynamics of nectar spurs, a key evolutionary innovation. *New Phytologist* 222: 1122–1138.

Ford KA. 2007. *Carex* (Cyperaceae)—Two new species from the calcareous mountains of North-West Nelson, New Zealand. *New Zealand Journal of Botany* 45: 721–730.

Ford BA, Ghazvini H, Naczi RFC, Starr JR. 2012. Phylogeny of *Carex* subg. *Vignea* (Cyperaceae) based on amplified fragment length polymorphism and nrDNA data. *Systematic Botany* 37: 913–925.

Gebauer S, Röser M, Hoffmann M. 2015. Molecular phylogeny of the species-rich *Carex* sect. *Racemosae* (Cyperaceae) based on four nuclear and chloroplast markers. *Systematic Botany* 40: 433–447.

Gebauer S, Starr JR, Hoffmann M. 2014. Parallel and convergent diversification in two northern hemispheric species-rich *Carex* lineages (Cyperaceae). *Organisms Diversity and Evolution* 14: 247–258.

Gehrke B, Linder HP. 2009. The scramble for Africa: Pan-temperate elements on the African high mountains. *Proceedings of the Royal Society B: Biological Sciences* 276: 2657–2665.

Gehrke B, Martín-Bravo S, Muasya A, Luceño M. 2010. Monophyly, phylogenetic position and the role of hybridization in *Schoenoxiphium* Nees (*Cariceae*, Cyperaceae). *Molecular Phylogenetics and Evolution* 56: 380–392.

Givnish TJ, Renner SS. 2004. Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *International Journal of Plant Sciences* 165 (Suppl 4): S1–S6.

Global Carex Group. 2015. Making *Carex* monophyletic (Cyperaceae, tribe *Cariceae*): A new broader circumscription. *Botanical Journal of the Linnean Society* 179: 1–42.

Global Carex Group. 2016a. Megaphylogenetic specimen-level approaches to the *Carex* (Cyperaceae) phylogeny using regions ITS, ETS, and *matK*: Implications for classification. *Systematic Botany* 41: 500–518.

Global Carex Group. 2016b. Specimens at the center: An informatics workflow and toolkit for specimen-level analysis of public DNA database data. *Systematic Botany* 41: 529–539.

Govaerts R, Jiménez-Mejías P, Koopman J, Simpson D, Goetghebeur P, Wilson K, Egorova T, Bruhl J. 2019. World Checklist of Cyperaceae. Facilitated by the Royal Botanic Gardens, Kew. Available from http://wcsp.science.kew.org/. [accessed May 2019].

Gradstein F, Ogg J, Smith A. 2004. *A geologic time scale*. Cambridge: Cambridge University Press.

Green PJ. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82: 711–732.

Hendrichs M, Michalski S, Begerow D, Oberwinkler F, Hellwig F. 2004a. Phylogenetic relationships in *Carex*, subgenus *Vignea* (Cyperaceae), based on ITS sequences. *Plant Systematics and Evolution* 246: 109–125.

Hendrichs M, Oberwinkler F, Begerow D, Bauer R. 2004b. *Carex*, subgenus *Carex* (Cyperaceae)—A phylogenetic approach using ITS sequences. *Plant Systematics and Evolution* 246: 89–107.

Hinchliff CE, Roalson EH. 2013. Using supermatrices for phylogenetic inquiry: An example using the sedges. *Systematic Biology* 62: 205–2019.

Hipp A, Rothrock P, Roalson E. 2009. The evolution of chromosome arrangements in *Carex* (Cyperaceae). *The Botanical Review* 75: 96–109.

Hipp AL. 2008. Phylogeny and patterns of convergence in *Carex* sect. *Ovales* (Cyperaceae): Evidence from ITS and 5.8S sequences. In: Naczi RFC, Ford BA eds. *Sedges: Uses, Diversity and Systematics of the Cyperaceae. Monographs in Systematic Botany from the Missouri Botanical Garden*. St. Louis: Missouri Botanic Garden Press. 108: 197–214.

Hipp AL, Manos PS, Hahn M, Avishai M, Bodénès C, Cavender-Bares J, Crowl A, Deng M, Denk T, Fitz-Gibbon, S, Gailing O, González-Elizondo MS, González-Rodríguez A, Grimm GW, Jiang XL, Kremer A, Lesur I, McVay JD, Plomion C, Rodríguez-Correa H, Schulze ED, Simeone MC, Sork VL, Valencia-Avalos S. 2019. The genomic landscape of the global oak phylogeny. *New Phytologist.* Available from https://doi.org/10.1101/587253

Hipp AL, Reznicek AA, Rothrock PE, Weber JA. 2006. Phylogeny and classification of *Carex* section *Ovales* (Cyperaceae). *International Journal of Plant Sciences* 167: 1029–1048.

Ho SYW, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology* 58: 367–380.

Hoffman MH, Gebauer S. 2016. Quantitative morphological and molecular divergence in replicated and parallel radiations in *Carex* (Cyperaceae) using symbolic data analysis. *Systematic Botany* 41: 552–557.

Hoffmann MH, Gebauer S, Rozycki TV. 2017. Assembly of the Arctic flora: Highly parallel and recurrent patterns in sedges (*Carex*). *American Journal of Botany* 104: 1–10.

Huang X, Deng T, Moore MJ, Wang H, Li Z, Lin N, Yusupov Z, Tojibaev KS, Wang Y, Sun H. 2019. Tropical Asian origin, boreotropical migration and long-distance dispersal in Nettles (Urticeae, Urticaceae). *Molecular Phylogenetics and Evolution* 137: 190–199.

Hug LA, Roger AJ. 2007. The impact of fossils and taxon sampling on ancient molecular dating analyses. *Molecular Biology and Evolution* 24: 1889–1897.

Hultén E. 1958. *The amphi-Atlantic plants and their phytogeographical connections*. Kungliga Svenska Vetenskapsakademien Handlingar, series 4, 7. Stockholm: Almquist and Wiksell.

Iglesias A, Artabe AE, Morel EM. 2011. The evolution of Patagonian climate and vegetation from the Mesozoic to the present. *Biological Journal of the Linnean Society* 103: 409–422.

Jiménez-Mejías P, Escudero M, Guerra-Cárdenas S, Lye KA, Luceño M. 2011. Taxonomic delimitation and drivers of speciation in the Ibero-North African *Carex* sect. *Phacocystis* river-shore group (Cyperaceae). *American Journal of Botany* 11: 1855–1867.

Jiménez-Mejías P, Luceño M, Amstein Lye K, Brochmann C, Gussarova G. 2012a. Genetically diverse but with surprisingly little geographical structure: The complex history of the widespread herb *Carex nigra* (Cyperaceae). *Journal of Biogeography* 39: 2279–2291.

Jiménez-Mejías P, Luceño M, Wilson KL, Waterway MJ, Roalson EH. 2016a. Clarification of the use of the terms perigynium and utricle in *Carex* L. (Cyperaceae). *Systematic Botany* 41: 519–528.

Jiménez-Mejías P, Martín-Bravo S, Luceño M. 2012b. Systematics and taxonomy of *Carex* sect. *Ceratocystis* (Cyperaceae) in Europe: A molecular and cytogenetic approach. *Systematic Botany* 37: 382–398.

Jiménez-Mejías P, Martinetto E. 2013. Toward an accurate taxonomic interpretation of *Carex* fossil fruits (Cyperaceae): A case study in section *Phacocystis* in the Western Palearctic. *American Journal of Botany* 100: 1580–1603.

Jiménez-Mejías P, Martinetto E, Momohara A, Popova S, Smith SY, Roalson EH. 2016b. A commented synopsis of the pre-Pleistocene fossil record of *Carex* (Cyperaceae). *The Botanical Review* 82: 258–345.

Johnson MA, Clark JR, Wagner WL, McDade LA. 2017. A molecular phylogeny of the Pacific clade of *Cyrtandra* (Gesneriaceae) reveals a Fijian origin, recent diversification, and the importance of founder events. *Molecular Phylogenetics and Evolution* 116: 30–48.

Kindlmann P, Schödelbauerová, I, Dixon A. 2007. Inverse latitudinal gradients in species diversity. In: Storch D, Marquet PA, Brown JH eds. *Scaling biodiversity*. Cambridge: Cambridge University Press. 246–257.

King MG, Roalson EH. 2009. Discordance between phylogenetics and coalescent-based divergence modelling: Exploring phylogeographic patterns of speciation in the *Carex macrocephala* species complex. *Molecular Ecology* 18: 468–482.

Koyama T. 1957. An enumeration of Hayata's Indo-Chinese collection of Cyperaceae. *Contributions de l'Institut Botanique de l'Université de Montréal* 70: 5–64.

Kükenthal G. 1909. Cyperaceae-Caricoideae. In: Engler HGA ed. *Das Pflanzenreich*. Leipzig: W. Engelmann. 4: 1–247.

Léveillé-Bourret É, Donadío S, Gilmour CN, Starr JR. 2015. *Rhodoscirpus* (Cyperaceae: Scirpeae), a new South American sedge genus supported by molecular, morphological, anatomical and embryological data. *Taxon* 64: 931–944.

Léveillé-Bourret É, Gilmour CN, Starr JR, Naczi RFC, Spalink D, Sytsma KJ. 2014. Searching for the sister to sedges (*Carex*): Resolving relationships in the Cariceae-Dulichieae-Scirpeae clade (Cyperaceae). *Botanical Journal of the Linnean Society* 176: 1–21.

Léveillé-Bourret É, Starr JR. 2019. Molecular and morphological data reveal three new tribes within the Scirpo-Caricoid Clade (Cyperoideae, Cyperaceae). *Taxon* 68: 218–245.

Léveillé-Bourret É, Starr JR, Ford BA. 2018a. A revision of *Sumatroscirpus* (Sumatroscirpeae, Cyperaceae) with discussions on Southeast Asian biogeography, general collecting, and homologues with *Carex* (Cariceae, Cyperaceae). *Systematic Botany* 43: 510–531.

Léveillé-Bourret É, Starr JR, Ford BA. 2018b. Why are there so many sedges? Sumatroscirpeae, a missing piece in the evolutionary puzzle of the giant genus *Carex* (Cyperaceae). *Molecular Phylogenetics and Evolution* 119: 93–104.

Léveillé-Bourret É, Starr JR, Ford BA, Moriarty Lemmon E, Lemmon AR. 2018c. Resolving rapid radiations within angiosperm families using anchored phylogenomics. *Systematic Biology* 67: 94–112.

Lewis AR, Marchant DR, Ashworth AC, Hedenäs L, Hemming SR, Johnson, JV, Leng MJ, Machlus ML, Newton AE, Raine JI, Willenbring JK, Williams M, Wolfe AP. 2008. Mid-Miocene cooling and the extinction of tundra in continental Antarctica. *Proceedings of the National Academy of Sciences USA* 105: 10676–10680.

Maguilla E, Escudero M, Luceño M. 2018. Vicariance versus dispersal across Beringian land bridges to explain circumpolar distribution: A case study in plants with high dispersal potential. *Journal of Biogeography* 45: 771–783.

Maguilla E, Escudero M, Waterway MJ, Hipp AL, Luceño M. 2015. Phylogeny, systematics, and trait evolution of *Carex* section *Glareosae*. *American Journal of Botany* 102: 1128–1144.

Mai DH, Walther H. 1988. Die pliozänen Floren von Thüringen, Deutsche Demokratische Republik. *Quartärpaläontologie* 7: 55–297.

Mairal M, Sanmartín I, Pellissier L. 2017. Lineage-specific climatic niche drives the tempo of vicariance in the Rand Flora. *Journal of Biogeography* 44: 911–923.

Manchester ST, Chen ZD, Lu AM, Uemura K. 2009. Eastern Asian endemic seed plant genera and their palaeogeographic history throughout the Northern Hemisphere. *Journal of Systematics and Evolution* 47: 1–42.

Márquez-Corro JI, Escudero M, Martín-Bravo S, Villaverde T, Luceño M. 2017. Long-distance dispersal explains the bipolar disjunction in *Carex macloviana*. *American Journal of Botany* 104: 663–673.

Márquez-Corro JI, Martín-Bravo S, Spalink D, Luceño M, Escudero M. 2019. Inferring hypothesis-based transitions in clade-specific models of chromosome number evolution in sedges (Cyperaceae). *Molecular Phylogenetics and Evolution* 135: 203–209.

Martín-Bravo S, Escudero M, Míguez M, Jiménez-Mejías, P, Luceño M. 2013. Molecular and morphological evidence for a new species from South Africa: *Carex rainbowii* (Cyperaceae). *South African Journal of Botany* 87: 85–91.

Matzke NJ. 2013. Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* 5: 242–248.

Matzke NJ. 2014a. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis with R Scripts, version 1.1.1. Available from https://rdrr.io/cran/BioGeoBEARS [accessed June 2019].

Matzke NJ. 2014b. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology* 63: 951–970.

Matzke NJ. 2016. Stochastic mapping under biogeographical models. PhyloWiki BiogeoBEARS website. Available from http://phylo.wikidot.com/biogeobears#stochastic_mapping [accessed June 2019]

Mello B, Schrago CG. 2014. Assignment of calibration information to deeper phylogenetic nodes is more effective in obtaining precise and accurate divergence time estimates. *Evolutionary Bioinformatics Online* 10: 79–85.

McGee D, Moreno-Chamarro E, Green B, Marshall J, Galbraith E, Bradtmiller L. 2018. Hemispherically asymmetric trade wind changes as signatures of past ITCZ shifts. *Quaternary Science Reviews* 180: 214–228.

Míguez M, Gehrke B, Maguilla E, Jiménez-Mejías P, Martín-Bravo S. 2017. *Carex* sect. *Rhynchocystis* (Cyperaceae): A Miocene subtropical relict in the Western Palaearctic showing a dispersal-derived Rand Flora pattern. *Journal of Biogeography* 44: 2211–2224.

Milne RI, Abbott RJ. 2002. The origin and evolution of Tertiary relict flora. *Advances in Botanical Research* 38: 281–314.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments (GCE10), 14 November, 2010. New Orleans, LA: IEEE. 1–8.

Mildenhall DC. 1980. New Zealand late Cretaceous and Cenozoic plant biogeography: A contribution. *Palaeogeography, Palaeoclimatology, Palaeoecology* 31: 197–233.

Molina A, Chung K-S, Hipp AL. 2015. Molecular and morphological perspectives on the circumscription of *Carex* section *Heleoglochin* (Cyperaceae). *Plant Systematics and Evolution* 301: 2419–2439.

Moreau CS, Bell CD. 2013. Testing the museum versus cradle tropical biological diversity hypothesis: Phylogeny, diversification, and ancestral biogeographic range evolution of the ants. *Evolution* 67: 2240–2257.

*J. Syst. Evol.* 57(6): 695–718, 2019

www.jse.ac.cn

Nelmes E. 1951. The genus *Carex* in Malaysia. *Reinwardtia* 1: 221–450.

Nogales M, Heleno R, Traveset A, Vargas P. 2012. Evidence for overlooked mechanisms of long-distance seed dispersal to and between oceanic islands. *New Phytologist* 194: 313–317.

Otero A, Jiménez-Mejías P, Valcárcel V, Vargas P. 2019. Being in the right place at the right time? Parallel diversification bursts favored by the persistence of ancient epizoochorous traits and hidden factors in Cynoglossoideae. *American Journal of Botany* 106: 438–452.

Pender JE. 2016. *Climatic niche estimation, trait evolution and species richness in North American Carex (Cyperaceae)*. PhD. Dissertation. Ottawa: University of Ottawa.

Plummer M, Best N, Cowles K, Vines K. 2006. CODA: Convergence diagnosis and output analysis for MCMC. *R News* 6: 7–11.

Pokorny L, Riina R, Mairal M, Meseguer AS, Culshaw V, Cendoya J, Serrano M, Carbajal R, Ortiz S, Heuertz M, Sanmartín I. 2015. Living on the edge: Timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. *Frontiers in Genetics* 6: 1–15.

Pole M. 2014. The Miocene climate in New Zealand: Estimates from paleobotanical data. *Palaeontologia Electronica* 17: 27A.

POWO. 2019. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Avalaible from http://www.plantsoftheworldonline.org [accessed 22 July 2019]

Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9: e89543

Rabosky DL, Grundler M, Anderson C, Title P, Shi JJ, Brown JW, Huang H, Larson JG. 2014. BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* 5: 701–707.

Raymond M. 1951. Sedges as material for phytogeographical studies. *Mémoires du Jardin Botanique de Montréal* 20: 2–23.

Raymond M. 1955. Cypéracées d'Indo-Chine. I. *Naturaliste Canadien* 82: 146–165.

Raymond M. 1959. Carices Indochinenses necnon Siamensis. *Mémoires du Jardin Botanique de Montréal* 53: 1–125.

Ree RH, Moore BR, Webb CO, Donoghue MJ. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59: 2299–2311.

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 Biogeography* 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.

Reznicek AA. 1990. Evolution in sedges (*Carex*, Cyperaceae). *Canadian Journal of Botany* 68: 1409–1432.

Roalson EH, Columbus JT, Friar EA. 2001. Phylogenetic relationships in Cariceae (Cyperaceae) based on ITS (nrDNA) and trnT-L-F (cpDNA) region sequences: Assessment of subgeneric and sectional relationships in *Carex* with emphasis on section *Acrocystis*. *Systematic Botany* 26: 318–341.

Roalson EH, Friar EA. 2004a. Phylogenetic relationships and biogeographic patterns in North American members of *Carex* section *Acrocystis* (Cyperaceae) using nrDNA ITS and ETS sequence data. *Plant Systematics and Evolution* 243: 175–187.

Roalson EH, Friar EA. 2004b. Phylogenetic analysis of the nuclear alcohol dehydrogenase (Adh) gene family in *Carex* section *Acrocystis* (Cyperaceae) and combined analyses of Adh and nuclear ribosomal ITS and ETS sequences for inferring species relationships. *Molecular Phylogenetics and Evolution* 33: 671–686.

Ronquist F. 1997. Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* 46: 195–203.

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.

Ruhfel BR, Bove CP, Philbrick CT, Davis CC. 2016. Dispersal largely explains the Gondwanan distribution of the ancient tropical clusioid plant clade. *American Journal of Botany* 103: 1117–1128.

Sanmartín I, Meseguer AS. 2016. Extinction in phylogenetics and biogeography: From timetrees to patterns of biotic assemblage. *Frontiers in Genetics* 7: 35.

Sanmartín I, Ronquist F. 2004. Southern hemisphere biogeography inferred by event-based models: Plant versus animal patterns. *Systematic Biology* 53: 216–243.

Sanmartín I, Wanntorp L, Winkworth RC. 2007. West wind drift revisited: Testing for directional dispersal in the Southern Hemisphere using event-based tree fitting. *Journal of Biogeography* 34: 398–416.

Saladin B, Leslie AB, Wüest RO, Litsios G, Conti E, Salamin N, Zimmermann NE. 2017. Fossils matter: Improved estimates of divergence times in *Pinus* reveal older diversification. *BMC Evolutionary Biology* 17: 95.

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

Sauquet H, Syw Ho, Gandolfo MA, Jordan GJ, Wilf P, Cantrill DJ, Bayly MJ, Bromham L, Brown GK, Carpenter RJ, Lee DM, Murphy DJ, Sniderman JM, Udovicic F. 2012. Testing the impact of calibration on molecular divergence times using a fossil-rich group: The case of *Nothofagus* (Fagales). *Systematic Biology* 61: 289–313.

Schellart WP, Lister G, Toy VG. 2006. A Late Cretaceous and Cenozoic reconstruction of the Southwest Pacific region: Tectonics controlled by subduction and slab rollback processes. *Earth-Science Reviews* 76: 191–233.

Schönberger I, Wilton AD, Boardman KF, Breitwieser I, Cochrane M, James de Lange P, de Pauw B, Fife AJ, Ford KA, Gibb ES, Glenny D, Korver M, Mosyakin SL, Novis PM, Prebble J, Redmond DN, Smissen RD, Tawiri K. 2017. *Checklist of the New Zealand flora—Seed plants*. Lincoln: Manaaki Whenua-Landcare Research.

Schönswetter P, Elven R, Brochmann C. 2008. Trans-Atlantic dispersal and large-scale lack of genetic structure in the circumpolar, arctic-alpine sedge *Carex bigelowii* s.l. (Cyperaceae). *American Journal of Botany* 95: 1006–1014.

Schönswetter P, Popp M, Brochmann C. 2006. Central Asian origin of and strong genetic differentiation among populations of the rare and disjunct *Carex atrofusca* (Cyperaceae) in the Alps. *Journal of Biogeography* 33: 948–956.

Semmouri I, Bauters K, Léveillé-Bourret È, Starr JR, Goetghebeur P, Larridon I. 2019. Phylogeny and systematics of Cyperaceae, the evolution and importance of embryo morphology. *Botanical Review* 85: 1–39.

Simões M, Breitkreuz L, Alvarado M, Baca S, Cooper JC, Heins L, Herzog K, Lieberman BS. 2016. The evolving theory of evolutionary radiations. *Trends in Ecology and Evolution* 31: 27–34.

Simpson MG, Johnson LA, Villaverde T, Guilliams CM. 2017. American amphitropical disjuncts: Perspectives from vascular plant

analyses and prospects for future research. *American Journal of Botany* 104: 1600–1650.

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

Smith SY, Collinson ME, Simpson DA, Rudall PJ, Marone F, Stampanoni M. 2009. Elucidating the affinities and habitat of ancient widespread Cyperaceae: *Volkeria messelensis* gen. et sp. nov., a fossil mapanioid sedge from the Eocene of Europe. *American Journal of Botany* 96: 1506–1518.

Spalink D, Drew BT, Pace MC, Zaborsky JG, Li P, Cameron KM, Givnish JG, Sytsma KJ. 2016a. Evolution of geographical place and niche space: Patterns of diversification in the North American sedge (Cyperaceae) flora. *Molecular Phylogenetics and Evolution* 95: 183–195.

Spalink D, Drew BT, Pace MC, Zaborsky JG, Starr JR, Cameron KM, Givnish TJ, Sytsma KJ. 2016b. Biogeography of the cosmopolitan sedges (Cyperaceae) and the area-richness correlation in plants. *Journal of Biogeography* 43: 1893–1904.

Spalink D, Pender J, Escudero M, Hipp AL, Roalson EH, Starr J, Waterway MJ, Bohs L, Sytsma KJ. 2018. The spatial structure of phylogenetic and functional diversity in the United States and Canada: An example using the sedge family (Cyperaceae). *Journal of Systematics and Evolution* 56: 449–465.

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

Starr JR, Bayer RJ, Ford BA. 1999. The phylogenetic position of *Carex* section *Phyllostachys* and its implications for phylogeny and subgeneric circumscription in *Carex* (Cyperaceae). *American Journal of Botany* 86: 563–577.

Starr JR, Ford BA. 2009. Phylogeny and evolution in Cariceae (Cyperaceae): current knowledge and future direction. *Botanical Review* 75: 110–137.

Starr JR, Harris SA, Simpson DA. 2004. Phylogeny of the unispicate taxa in Cyperaceae tribe Cariceae I: Generic relationships and evolutionary scenarios. *Systematic Botany* 29: 528–544.

Starr JR, Harris SA, Simpson DA. 2008. Phylogeny of the unispicate taxa in Cyperaceae tribe Cariceae II: the limits of *Uncinia*. In: Naczi RFC, Ford BA eds. *Sedges: Uses, Diversity and Systematics of the Cyperaceae. Monographs in Systematic Botany from the Missouri Botanical Garden* 108: 243–267.

Starr JR, Janzen FH, Ford BA. 2015. Three new early diverging *Carex* (Cariceae, Cyperaceae) lineages from East and Southeast Asia with important evolutionary and biogeographic implications. *Molecular Phylogenetics and Evolution* 88: 105–120.

Svenning C. 2003. Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters* 6: 646–653.

Stehli FG, Webb D. 1985. *The great American interchange*. New York: Plenum Press.

Takhtajan A. 1986. *Floristic regions of the world*. Berkeley: University of California Press.

Thiers B. 2019. Index Herbariorum: A Global Directory of Public Herbaria and Associated Staff. New York Botanical Garden's Virtual Herbarium. Available from http://sweetgum.nybg.org/science/ih/ [accessed November 2019].

Tripp EA, McDade LA. 2014. A rich fossil record yields calibrated phylogeny for Acanthaceae (Lamiales) and evidence for marked biases in timing and directionality of intercontinental disjunctions. *Systematic Biology* 63: 660–684.

Uribe-Convers S, Tank DC. 2015. Shifts in diversification rates linked to biogeographic movement into new areas: An example of

disparate continental distributions and a recent radiation in the Andes. *American Journal of Botany* 102: 1854–1869.

Uzma, Jiménez-Mejías P, Amir R, Qasim Hayat M, Hipp AL. 2019. Time and ecological priority shaped the diversification of sedges in the Himalayas. *PeerJ* 7: e6792

Vargas P, Valente LM, Blanco-Pastor JL, Liberal I, Guzmán B, Cano E, Forrest A, Fernández-Mazuecos M. 2014. Testing the biogeographical congruence of palaeofloras using molecular phylogenetics: Snapdragons and the Madrean-Tethyan flora. *Journal of Biogeography* 41: 932–943.

Veevers JJ, Powel CM, Roots SR. 1991. Review of the seafloor spreading around Australia. I. Synthesis of the patterns of spreading. *Australian Journal of Earth Sciences* 38: 373–389.

Viljoen J, Muasya MA, Barrett RL, Bruhl JJ, Gibbs AK, Slingsby JA, Wilson KA, Verboom GA. 2013. Radiation and repeated transoceanic dispersal of Schoeneae (Cyperaceae) through the southern hemisphere. *American Journal of Botany* 100: 2494–2508.

Villaverde T, Escudero M, Luceño M, Martín-Bravo S. 2015a. Long-distance dispersal during the middle–late Pleistocene explains the bipolar disjunction of *Carex maritima* (Cyperaceae). *Journal of Biogeography* 42: 1820–1831.

Villaverde T, Escudero M, Martín-Bravo S, Bruederle LP, Luceño M, Starr JR. 2015b. Direct long-distance dispersal best explains the bipolar distribution of *Carex arctogena* (*Carex* sect. *Capituligerae*, Cyperaceae). *Journal of Biogeography* 42: 1514–1525.

Villaverde T, Escudero M, Martín-Bravo S, Jiménez-Mejías P, Sanmartín I, Vargas P, Luceño M. 2017a. Bipolar distributions in vascular plants: A review. *American Journal of Botany* 104: 1680–1694.

Villaverde T, Escudero M, Martín-Bravo S, Luceño M. 2017b. Two independent dispersals to the Southern Hemisphere to become the most widespread bipolar *Carex* species: Biogeography of *C. canescens* (Cyperaceae). *Botanical Journal of the Linnean Society* 183: 360–372.

Villaverde T, Jiménez-Mejías P, Luceño M, Waterway MJ, Kim S, Lee B, Rincón-Barrado M, Hahn M, Maguilla E, Roalson EH, Hipp AL, Global Carex Group. A new classification of *Carex* subgenera supported by a HybSeq backbone phylogeny. [In review]

Villaverde T, Maguilla E, Escudero M, Márquez-Corro JI, Jiménez-Mejías P, Gehrke B, Martín-Bravo S, Luceño M. 2017c. New insights into the systematics of the Schoenoxiphium clade (*Carex*, Cyperaceae). *International Journal of Plant Sciences* 178: 320–329.

Waterway MJ, Hoshino T, Masaki T. 2009. Phylogeny, species richness, and ecological specialization in Cyperaceae tribe Cariceae. *Botanical Review* 75: 138–159.

Waterway MJ, Starr JR. 2007. Phylogenetic relationships in tribe Cariceae (Cyperaceae) based on nested analyses of four molecular data sets. *Aliso* 23: 165–192.

WCSP. 2019. World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. Avalaible from http://wcsp.science.kew.org/ [accessed 22 July 2019].

Wellborn GA, Langerhans RB. 2015. Ecological opportunity and the adaptive diversification of lineages. *Ecology and Evolution* 5: 176–195.

Westergaard KB, Zemp N, Bruederle LP, Stenøien HK, Widmer A, Fior S. 2019. Population genomic evidence for plant glacial survival in Scandinavia. *Molecular Ecology* 28: 818–832.

Willis CG, Davis CC. 2015. Rethinking migration. *Science* 348: 766.

Willis CG, Franzone BF, Xi Z, Davis CC. 2014. The establishment of Central American migratory corridors and the biogeographic

origins of seasonally dry tropical forests in Mexico. *Frontiers in Genetics* 5: 433.

Yano O, Ikeda H, Jin XF, Hoshino T. 2014. Phylogeny and chromosomal variations in East Asian *Carex*, *Siderosticate* group (Cyperaceae), based on DNA sequences and cytological data. *Journal of Plant Research* 127: 99–107.

Zhisheng A, Kutzbach JE, Prell WL, Porter SC. 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. *Nature* 411: 62–66.

Zuloaga FO, Salariato DL, Scataglini A. 2018. Molecular phylogeny of *Panicum* s. str. (Poaceae, Panicoideae, Paniceae) and insights into its biogeography and evolution. *PLoS ONE* 13: e0191529

# Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.12549/suppinfo:

**Data S1.** List of taxa sampled in the present study (species, subspecies, and varieties) and geographical ascription to the six areas coded in the biogeographic analyses (see Materials & Methods; Fig. 1B). Subgeneric placement follows Villaverde et al. (in review). Sectional placement has been modified from GCG (2016a), and must be considered merely indicative. Geographical distribution has been coded according to Govaerts et al. (2019).

**Data S2.** Material studied. Specimens sequenced by GCG (2016a) and the present study are listed in one spreadsheet and labeled with a specimen number (spm). Material downloaded from GenBank published by GCG (2016b) and subsequent studies is displayed in a different spreadsheet and labeled with "NCBI" followed by a voucher code. Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Data S3.** Unconstrained multiple tips query tree inferred from the multiple tips matrix, built under the GCG (2016a) scaffolding approach using ML and the evolutionary placement algorithm as implemented in RAxML. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001). Supports are SH-aLRT values. Well-supported branches must be considered when SH-aLRT values are above 80%.

**Data S4.** Unconstrained singletons query tree built from the singletons matrix, built under the GCG (2016a) scaffolding approach using ML and the evolutionary placement algorithm as implemented in RAxML. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001). Supports are SH-aLRT values. Well-supported branches must be considered when SH-aLRT values are above 80%.

**Data S5.** ML constrained singletons tree, built from the singletons matrix and constrained using Villaverde et al. (in review) phylogenomic topology. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001). Branch supports are bootstrap values.

**Data S6.** Chronogram based on the constrained singletons query tree (Data S5) obtained with TreePL, using ten fossils placed both at deep and shallow nodes as calibration constraints (Table 1). Ages are indicated at nodes. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Data S7.** Chronogram based on the constrained singletons tree (Data S5) obtained with TreePL, using three deep-node fossils as calibration constraints (Table 1). Ages are indicated at nodes. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Data S8.** Chronogram based on the unconstrained singletons query tree (Data S4) obtained with TreePL, using ten fossils placed both at deep and shallow nodes as calibration constraints (Table 1). Ages are indicated at nodes. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Data S9.** Chronogram based on the unconstrained singletons query tree (Data S4) obtained with TreePL, using three deep-node fossils as calibration constraints (Table 1). Ages are indicated at nodes. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Fig. S1.** Ancestral area reconstruction under DEC model as implemented in BioGeoBEARS, based on the constrained singletons chronogram calibrated using ten fossils (Data S6). The first tree depicts the most probable area or combination of areas with colored squares on each node. The second tree shows the probability of the different competing areas with colored circle pies on each node. Color of squares and circle pies according to Fig. 1B. Letters at nodes refer to the biogeographical region inferred: N: Nearctic; E: W Palearctic; A: E Asia; F: Afrotropic; S: Neotropic; P: Pacific. Tip labels indicate whether the corresponding specimen was sequenced

www.jse.ac.cn

*J. Syst. Evol.* 57(6): 695–718, 2019

by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Fig. S2.** Ancestral area reconstruction under DEC+J model as implemented in BioGeoBEARS, based on the constrained singletons chronogram calibrated using ten fossils (Data S6). The first tree depicts the most probable area or combination of areas with colored squares on each node. The second tree shows the probability of the different competing areas on each node with colored circle pies. Color of squares and circle pies according to Fig. 1B. Letters at nodes refer to the biogeographical region inferred: N: Nearctic; E: W Palearctic; A: E Asia; F: Afrotropic; S: Neotropic; P: Pacific. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Fig. S3.** Ancestral area reconstruction under DIVA model as implemented in BioGeoBEARS, based on the constrained singletons chronogram calibrated using ten fossils (Data S6). The first tree depicts the most probable area or combination of areas with colored squares on each node. The second tree shows the probability of the different competing areas on each node with colored circle pies. Color of squares and circle pies according to Fig. 1B. Letters at nodes refer to the biogeographical region inferred: N: Nearctic; E: W Palearctic; A: E Asia; F: Afrotropic; S: Neotropic; P: Pacific. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the

specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Fig. S4.** Ancestral area reconstruction under DIVA+J model as implemented in BioGeoBEARS, based on the constrained singletons chronogram calibrated using ten fossils (Data S6). The first tree depicts the most probable area or combination of areas with colored squares on each node. The second tree shows the probability of the different competing areas on each node node with colored circle pies. Color of squares and circle pies according to Fig. 1B. Letters at nodes refer to the biogeographical region inferred: N: Nearctic; E: W Palearctic; A: E Asia; F: Afrotropic; S: Neotropic; P: Pacific. Tip labels indicate whether the corresponding specimen was sequenced by GCG (2016a) or the present study (tips labeled with a specimen (spm) number), or was obtained from GenBank (tips labeled with "NCBI" followed by a voucher code). Tip labels also include the geographical origin of the specimen using TDWG level 3 regions abbreviations ("botanical countries"; Brummitt, 2001).

**Fig. S5.** Timing of the different kind of biogeographic events inferred by the BSM analysis under DEC model.

**Table S1. A,** Summary of the number of dispersal events (and standard deviations) between the different considered regions (sources of dispersal in rows and sinks of dispersals in columns) inferred by the BSM analysis under the DEC+J model. **B,** Summary of the number of cladogenetic dispersal (founder) events (and standard deviations) between the different considered regions (sources of dispersal in rows and sinks of dispersals in columns) inferred by the BSM analysis under the DEC+J model. **C,** Summary of the number of anagenetic dispersal (range expansion) events (and standard deviations) between the different considered regions (sources of dispersal in rows and sinks of dispersals in columns) inferred by the BSM analysis under the DEC+J model. Cell color indicates the range of the number of inferred dispersal events, as indicated below each table. **D,** Summary of the type of biogeographical events inferred with BSM under the DEC+J model.