Botanical Journal of the Linnean Society, 2020, **194**, 141–163. With 4 figures.

A new classification of *Carex* (Cyperaceae) subgenera supported by a HybSeq backbone phylogenetic tree

TAMARA VILLAVERDE[1,2,3,*,†], PEDRO JIMÉNEZ-MEJÍAS[4,5,6,*], MODESTO LUCEÑO[2], MARCIA J. WATERWAY[7], SANGTAE KIM[8], BORA LEE[8], MARIO RINCÓN-BARRADO[3], MARLENE HAHN[1], ENRIQUE MAGUILLA[9], ERIC H. ROALSON[6] and ANDREW L. HIPP[1,10,*]; THE GLOBAL *CAREX* GROUP[#]

[1]*The Morton Arboretum, 4100 Illinois Route 53, Lisle IL 60532, USA*
[2]*Área de Botánica, Department of Molecular Biology and Biochemical Engineering, Universidad Pablo de Olavide, ctra de Utrera km 1 sn, 41013 Seville, Spain*
[3]*Real Jardín Botánico, CSIC, Plaza de Murillo 2, 28014 Madrid, Spain*
[4]*Departamiento de Biología (Botánica), Facultad de Ciencias Biológicas, Universidad Autónoma de Madrid, c/ Darwin 2, 28049 Madrid, Spain*
[5]*Centro de Investigación en Biodiversidad y Cambio Global (CIBC-UAM), Universidad Autónoma de Madrid, 28049 Madrid, Spain*
[6]*School of Biological Sciences, Washington State University, 99164 Pullman WA, USA*
[7]*Plant Science Department, McGill University, Ste-Anne-de-Bellevue, H9X 3V9, Quebec, Canada*
[8]*Department of Biology, Sungshin University, Seoul 01133, Korea*
[9]*Department of Plant Biology and Ecology, University of Seville, Reina Mercedes sn, ES-41012 Seville, Spain*
[10]*The Field Museum, 1400 S Lake Shore Drive, Chicago IL 60605, USA*
[†]*Current address: Departamento de Biología y Geología, Universidad de Almería, Carretera de Sacramento s/n, La Cañada de San Urbano, 04120, Almería, Spain*

Received 29 August 2019; revised 23 March 2020; accepted for publication 29 May 2020

The field of systematics is experiencing a new molecular revolution driven by the increased availability of high-throughput sequencing technologies. As these techniques become more affordable, the increased genomic resources have increasingly far-reaching implications for our understanding of the Tree of Life. With *c*. 2000 species, *Carex* (Cyperaceae) is one of the five largest genera of angiosperms and one of the two largest among monocots, but the phylogenetic relationships between the main lineages are still poorly understood. We designed a Cyperaceae-specific HybSeq bait kit using transcriptomic data of *Carex siderosticta* and *Cyperus papyrus*. We identified 554 low-copy nuclear orthologous loci, targeting a total length of *c.* 1 Mbp. Our Cyperaceae-specific kit shared loci with a recently published angiosperm-specific Anchored Hybrid Enrichment kit, which enabled us to include and compile data from different sources. We used our Cyperaceae kit to sequence 88 *Carex* spp., including samples of all the five major clades in the genus. For the first time, we present a phylogenetic tree of *Carex* based on hundreds of loci (308 nuclear exon matrices, 543 nuclear intron matrices and 66 plastid exon matrices), demonstrating that there are six strongly supported main lineages in *Carex*: the Siderostictae, Schoenoxiphium, Unispicate, Uncinia, Vignea and Core Carex clades. Based on our results, we suggest a revised subgeneric treatment and provide lists of the species belonging to each of the subgenera. Our results will inform future biogeographic, taxonomic, molecular

*Corresponding author. E-mails: tvilhid@gmail.com; pjimmej@gmail.com; ahipp@mortonarb.org.
[#]The Global *Carex* Group have contributed to the writing and assembly of this paper in various ways and through formal and informal discussions.

dating and evolutionary studies in *Carex* and provide the step towards a revised classification that seems likely to stand the test of time.

ADDITIONAL KEYWORDS: Cyperaceae – genomics – hyperdiverse – nomenclature – systematics – subgenera – targeted sequencing.

INTRODUCTION

Molecular systematics has reshaped our understanding of the Tree of Life at all taxonomic levels, from subspecies to domains (Vargas & Zardoya, 2014). One of the great successes has been the revised classification of angiosperms into classes and families that reflect phylogeny (APG IV, 2016, and earlier APG accounts). Within families, the process of reorganizing angiosperm taxonomy tends to advance more slowly for the most species-rich clades because of their complexity and the large amounts of data that need to be collected. Revisions of large families may take decades and many rounds of revisions (e.g. Apiaceae: Hardway *et al.*, 2004; Nicolas & Plunkett, 2009; Downie *et al.*, 2010; Jiménez-Mejías & Vargas, 2015; Banasiak *et al.*, 2016; Asteraceae, Cichorieae: Kilian, Hand & Raab-Straube 2009+; Fabaceae: Käss & Wink 1996; Allan & Porter 2000; Cardoso *et al.*, 2012; LPWG, 2017; de la Estrella *et al.*, 2018; Poaceae: Soreng *et al.*, 2017, among many others). The classification of hyperdiverse genera (e.g. *Euphorbia* L., with *c*. 2000 species; Yang *et al.*, 2012; Dorsey *et al.*, 2013; Peirson *et al.*, 2013; Riina *et al.*, 2013) has been particularly difficult in terms of species sampling, assessing homoplasy and testing relationships with satellite genera.

With *c*. 2000 species, *Carex* L. (Cyperaceae) is one of the five largest genera of angiosperms and one of the two largest in monocots, with the orchid genus *Bulbophyllum* Thouars (WCSP, 2019). *Carex* is placed in tribe Cariceae, which are characterized by having unisexual flowers, the female ones contained within a prophyllar structure called a perigynium, which is referred to as a utricle when its margins are fused and closed (Jiménez-Mejías *et al.*, 2016). The first phylogenetic analyses of the genus (Starr, Bayer & Ford, 1999; Yen & Olmstead, 2000; Roalson, Columbus & Friar, 2001;Hendrichs *et al.*, 2004a, b; Starr, Harris & Simpson 2004) made clear that generic delimitations in Cariceae and the infrageneric classification of *Carex* were unnatural. A first major accomplishment was the recognition that the smaller genera previously recognized in Cariceae (*Cymophyllus* Mack., *Kobresia* Willd., *Schoenoxiphium* Nees and *Uncinia* Pers.) are nested in *Carex*. Consequently, the Global Carex Group (2015) transferred all genera and species of Cariceae into *Carex* (Cariceae = *Carex* hereafter). A second major discovery was that most traditionally recognized subgenera of *Carex* are polyphyletic (Global Carex Group 2016a, b, and references therein). However, a revised subgeneric classification of *Carex* has not yet been proposed.

In *Carex*, as in other large genera, species identification is greatly facilitated by infrageneric classification (Global Carex Group, 2016a). The four traditional subgenera of *Carex* have served as a classification gateway to the numerous sections and species. These subgenera, until recently widely accepted by the community of cyperologists, are largely based on the work of Kükenthal (1909), who defined them according to inflorescence structure, sex distribution of flowers and presence and morphology of cladoprophylls (small sheathing bracts associated to the distal branches the sedge inflorescences). *Carex* subgenus *Primocarex* Kük. contained all species with inflorescence reduced to a single terminal spike. *Carex* subgenus *Vignea* (P.Beauv ex T.Lestib.) Peterm. was composed of species without cladoprophylls and with predominantly sessile bisexual spikes. *Carex* subgenus *Indocarex* (Baill.) Kük. included species with utriculiform cladoprophylls and bisexual (androgynous) spikes. Finally, *Carex* subgenus *Carex* (as *Eucarex* Peterm. in Kükenthal 1909) was the largest subgenus and comprised a diverse conglomerate of species mostly characterized by having a tubular cladoprophyll (Fig. 1; prophyll terminology according to Jiménez-Mejías *et al.*, 2016). Since Kükenthal (1909), the only departure from this scheme has been Egorova's (1999) revision of *Carex* of the former USSR, which updated the nomenclature of the four subgenera proposed by Kükenthal. *Vignea* remained unchanged, *Carex* (=*Eucarex*), *Psyllophorae* (Degl.) Peterm. (= *Primocarex*) and *Vigneastra* (Tuck.) Kük (=*Indocarex*) were accepted as updated names, and a fifth subgenus, *Carex* subgenus *Kreczetoviczia* T.V.Egorova, was added to accommodate most of the species with two stigmas traditionally placed in subgenus *Carex*. Most recent floristic treatments that use a subgeneric classification for *Carex* have relied roughly on Kükenthal's subgenera (e.g. Chater, 1980; Kukkonen, 1996; Ball & Reznicek, 2002; Luceño, Escudero & Jiménez-Mejías, 2008), perhaps with the single exception of the treatment by Dai *et al.*, (2010) for the *Flora of China*, in which only three subgenera were recognized (*Vignea*, *Vigneastra* and *Carex*).

Although Kükenthal's subgenera provided a convenient way to divide *Carex* into smaller groups, his classification was called into question almost

Figure 1. Inflorescence diversity in the subgenera recognized in the present study. A, Siderostictae clade, *Carex sidérosticta*. B–D, Schoenoxiphium clade (B, C. *baldensis*; C, *C. parvirufa*; D, *C. camptoglochin*). E–G, Unispicate clade (E, *C. microglochin*; F, *C. nardina*; G, *C. simpliciuscula*). H, I, Uncinia clade (H, *C. lechleriana*; I, *C. cordillerana*). J–L, Vignea clade (J, *C. canescens*; K, *C. dioica*; L, *C. maritima*). M–P, Core Carex clade (M, *C. macrosolen*; N, *C. alba*; O, *C. pseudocyperus*; P, *C. wahlenbergiana*).

20 years ago with the first molecular studies of the genus (Starr *et al.*, 1999; Yen & Olmstead, 2000; Roalson *et al.*, 2001; Hendrichs *et al.*, 2004a, b; Starr *et al.*, 2004). In a notable landmark paper, Waterway & Starr (2007) used four nuclear ribosomal and plastid loci on an extensive species sampling across the genus to show that *Carex* comprised four major clades: (1) the Core Carex clade, containing most species belonging to subgenus *Carex* plus subgenus *Vigneastra*; (2) the Vignea clade, comprising species of subgenus *Vignea*; (3) the Schoenoxiphium clade, which grouped a few species of subgenus *Psyllophorae* with species of the former genus *Schoenoxiphium* and (4) the 'Core' Unispicate clade, bringing together the rest of species from subgenus *Psyllophorae*, species of the genera *Uncinia* and *Kobresia* and *Carex curvula* All., the latter formerly placed in subgenus *Vignea*. However, the authors were unable to resolve relationships among these four lineages. Although the Unispicate and Schoenoxiphium clades were resolved as sister to each other with strong support (forming the so-called Caricoid clade), the branching order among the Core Carex and Vignea clades lacked support. Waterway, Hoshino & Masaki (2009), using the same four markers, revealed a fifth early-diverging major lineage, (5) the Siderostictae clade, comprising the species of *Carex* section *Siderostictae* Franch. ex Ohwi. Although this group was strongly supported as sister to the rest of the genus, the internal relationships among the Caricoid, Vignea and Core Carex clades remained unresolved. Subsequent phylogenetic works, adding new species to the sampling, further clarified which species and sections comprised each of the clades (e.g. Siderostictae clade: Yano *et al.*, 2014; Siderostictae and Carex clades: Starr, Janzen & Ford, 2015; Schoenoxiphium clade: Gerkhe *et al.*, 2010, Villaverde *et al.*, 2017; *Carex* as a whole: Waterway *et al.*, 2015) but at the same time unmasked new conflicts: Starr *et al.* (2015), for example, found that the Caricoid clade was only resolved when using likelihood and Bayesian inference and for certain sets of sampled genes. A substantial increase in taxon sampling was published soon thereafter by the Global Carex Group (2016a), who inferred a phylogenetic tree with 996 species (*c.* 50% of the genus) representing 110 of the 126 recognized sections (*c.* 92%) based on three loci (nuclear ITS and ETS and plastid *matK*). They did not retrieve a Caricoid clade, although the phylogenetic hypothesis still yielded the five major clades mentioned above. A recent phylogenetic analysis based on anchored phylogenomics (461 nuclear loci;Léveillé-Bourret *et al.*, 2018a) and a limited sampling of *Carex* spp. (18) confirmed the Siderostictae clade as sister to the rest of the genus, although the branching order among the major clades again showed considerable uncertainty. The structure of the major *Carex* phylogenetic hypotheses produced to date are summarized in Table 1.

The field of systematics is experiencing a new molecular revolution driven by the rapidly increasing use of phylogenomic approaches (e.g. Harrison & Kidner, 2011; Lemmon & Lemmon, 2013; Vargas, Ortiz & Simpson, 2017). The variety of approaches for economically sequencing hundreds to thousands of loci from across the genomes of hundreds of individuals has enabled systematists to investigate species boundaries and higher-level relationships with unprecedented precision. Among the various techniques, the combination of targeted sequencing with genome skimming, or the HybSeq approach, has emerged as a fast and cost-efficient method for sequencing hundreds of single-copy nuclear loci across numerous individuals for addressing phylogenetic and population genetic questions, usually yielding highly supported evolutionary relationships (Weitemier *et al.*, 2014; Kates *et al.*, 2018; Villaverde *et al.*, 2018). The method has proven useful even with old material conserved in biological collections ('museomics'; Bakker, 2017). The HybSeq method targets orthologous loci identified using reference genomes and transcriptomes of more or less closely related organisms, using baits designed to capture these conserved loci and their flanking regions (Lemmon, Emme & Lemmon, 2012; Johnson *et al.*, 2016; Kates *et al.*, 2018). Moreover, the genome skimming that occurs during the high-throughput sequencing usually recovers some off-target plastome or nuclear ribosomal data (e.g. Kriebel *et al.*, 2019), which may also be used to explore phylogenetic relationships.

In this study, our goal was to infer phylogenetic relationships among the main lineages in *Carex* using a new set of sequence baits designed to capture > 500 genes in Cyperaceae. Because they were developed for Cyperaceae, which diversified *c.* 77–89 Mya (Spalink *et al.*, 2016), rather than the angiosperms as a whole, which diversified *c.* 140–250 Mya (Sauquet & Magallón, 2018), we expected that our markers would yield more variable regions and higher sequence capture efficiency than a 'one-size-fits-all' set (Kadlec *et al.*, 2017), such as those presented recently for anchored phylogenomics (Léveillé-Bourret *et al.*, 2018a) or targeted sequencing using angiosperm-wide baits (Johnson *et al.*, 2018). In particular, we aimed to (1) test the phylogenetic hypothesis that *Carex* comprises five major clades, as reported to date; and (2) reconstruct evolutionary relationships among them using nuclear and plastid data. Based on our results, we propose a revised subgeneric treatment and enumerate the species belonging to each of the subgenera.

Table 1. Comparison of the tree topologies from combined datasets obtained by the major Sanger-based phylogenetic works focusing on *Carex* by chronological order

Publication	Markers	Siderostictae clade	Vignea clade	Core Carex clade	Caricoid clade	Schoenoxiphium clade	Unispicate clade	Uncinia clade
Starr *et al.* (1999)	ITS	-	-	MP: 85	-	-	MP: 81	
Yen & Olmstead (2000)	*ndhF*, *trnL*, *trnL-trnF*	-	MP: 100 **ML**	MP: 100 **ML**	-	MP: 100 **ML**	MP: not retrieved ML	
Roalson *et al.* (2001)	ITS, *trnL*, *trnL-trnF*	-	MP: 100	MP: 97	MP: not found	Only one sample included	MP: not found	
Starr, Harris & Simpson (2004,2008)	ETS, ITS	-	**ML**	**ML**	**ML**	**ML**	**ML**	
Waterway & Starr (2007)	ETS, ITS, *trnE-trnD*, *trnL*, *trnL-trnF*,	-	MP: 100 BI: 100	MP: 99 BI: 100	MP: not found BI: 98	MP: 73 BI: 100	MP: 59 BI: 100	
Waterway *et al.* (2009)	ETS, ITS, *trnL*, *trnL-trnF*,	MP: 100 BI: > 95	MP: 100 BI: > 95	MP: 100 BI: > 95	MP: not found BI: 91–95	MP: 59 BI: > 95	MP: not found BI: > 95	
Starr & Ford (2009)	ETS, ITS	-	MP: 98 BI: 100	MP: 97 BI: 100	MP: not found BI: 95	MP: 73 BI: 100	MP: not found BI: 100	
Gehrke *et al.*, (2010)	ITS, *rps16*, *trnL-F*	-	MP: 85 BI: 100	MP: 100 BI: 100	MP: not found BI: not found	MP: 99 BI: 100	MP: 54 BI: 100	
Starr *et al.* (2015)	ETS, ITS, *matK*, *ndhF*, *rps16*	MP: 90 BI: > 95	MP: 100 BI: > 95	MP: 99 BI: > 95	MP: not found BI: < 95	MP: 100 BI: > 95	MP: 93 BI: > 95	
Global Carex Group (2016a)	ETS, ITS, *matK*	ML: 91	ML: 97	ML: 100	ML: not found	ML: 98	ML: 97	
This study (nuclear exons)	308 loci	ML:100 LPP:1	ML:100 LPP:1	ML:100 LPP:1	ML: not found	ML:100 LPP:1	ML:100 LPP:1	ML:100 LPP:0.96
This study (plastid exons)	66 loci	ML:100 LPP:1	ML:100 LPP < 0.9	ML:100 LPP:1	ML:96	ML:100 LPP:1	ML:100 LPP < 0.9	ML:93 LPP < 0.9

For the clades recovered in each phylogeny we provide the support retrieved. Analyses are abbreviated as follow: Bayesian inference (BI), local posterior probability (LPP), maximum likelihood (ML) and maximum parsimony (MP). ML in bold denotes that this means that the clade was recovered but without significant support.

MATERIAL AND METHODS

TAXON SAMPLING

We sampled 88 *Carex* spp. using herbarium and freshly collected silica-dried samples of all five major clades of the genus (Siderostictae, Schoenoxiphium, Unispicate, Vignea and Core Carex clades; see Supporting Information, Table S1). Our study covers the phylogenetic diversity of the genus proportionally as demonstrated in the previous (*c.* 1000-species) phylogenetic analysis of the Global Carex Group (2016a). The existence of the Global *Carex* Group as an international collaboration for the study of the systematics of *Carex* made it possible for us to sample optimally from across the phylogenetic tree. We incorporated sequenced samples (ten and 33 samples for the nuclear and plastid analyses, respectively) from Léveillé-Bourret *et al.* (2018a) to test whether the

gene coverage of their angiosperm-specific Anchored Hybrid Enrichment kit (Buddenhagen *et al.*, 2016) would integrate well with the genes targeted by our Cyperaceae-specific kit (see Supporting Information, Tables S1, S2). We also used four and 18 samples belonging to other genera of Cyperaceae (Supporting Information, Tables S1, S2) in the nuclear and plastid phylogenetic reconstructions, respectively (see below), as additional species in the outgroup. We did not include samples from *Sumatroscirpus* Oteng-Yeboah, the sister-genus of *Carex* (Léveillé-Bourret, Starr & Ford, 2018b), as it was unknown at the beginning of this study.

Transcriptomic data

Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Healden, Germany) from a single plant of *Carex siderosticta* Hance (voucher: *S. Kim 2010001*, SWU) cultivated in the greenhouse of Sungshin University. Total RNA samples were converted into cDNA libraries using the TruSeq Stranded mRNA Sample Prep Kit (Illumina). Starting with 1000 ng of total RNA, poly-adenylated RNA (primarily mRNA) was selected and purified using oligo-dT-conjugated magnetic beads. This mRNA was physically fragmented and converted into single-stranded cDNA using reverse transcriptase and random hexamer primers, with the addition of actinomycin D to suppress DNA-dependent synthesis of the second strand. Double-stranded cDNA was created by removing the RNA template and synthesizing the second strand in the presence of dUTP in place of dTTP. A single A base was added to the 3′ end to facilitate ligation of sequencing adapters, which contain a single T base overhang. Adapter-ligated cDNA was amplified by polymerase chain reaction to increase the amount of sequence-ready library. During this amplification, the polymerase stalls when it encounters a U base, rendering the second strand a poor template. Accordingly, amplified material used the first strand as a template, thereby preserving the strand information. Final cDNA libraries were analysed for size distribution and using an Agilent Bioanalyzer (DNA 1000 kit; Agilent), quantitated by qPCR (Kapa Library Quant Kit; Kapa Biosystems, Wilmington, MA), then normalized to 2 nmol/L in preparation for sequencing. Library was sequenced using the Illumina HiSeq2000 platform by Macrogen Co. Ltd (Seoul).

A total of 52 937 818, 2 × 100-bp paired-end reads (5.34 Gbp) were generated, of which 52 499 568 reads (5.23 Gbp) remained after quality-filtration using Trimmomatic v.0.33 (Bolger, Lohse & Usadel, 2014) with the default option: phred score 33, sliding window size 4, cutoff quality 15, leading base 3, trailing base 3, crop length 0, headcrop length 0 and minimum length 36.

De novo assembly was performed using SOAPdenovo-Trans v.1.03 (Xie *et al.*, 2014) with the options: k-mer 25, merge level 3, with scaffold and fillgap option activated. Based on high-quality reads, 72 279 transcripts were assembled (47 778 184 bp) with an N50 of 1697 bp. The average length of assembled transcripts was 675 bp.

Bait design and sequence capture

Baits were designed from transcriptomic data from *Carex siderosticta* and a relatively distantly related species of Cyperaceae, *Cyperus papyrus* L. (available through the 1KP initiative: www.onekp.com/public_data.html), using the MarkerMiner v.1 pipeline (Chamala *et al.*, 2015). We identified orthologous low-copy nuclear genes (LCNGs) using the proteome of *Oryza sativa* L. as a reference. Five hundred and fifty-four orthologous loci were selected to develop the gene target baits, which ranged from 802 to 7445 bp, for a total exon length of 1 032 784 bp (Supporting Information, Table S3). The target enrichment kit was manufactured as 15 075 120-bp baits with in-solution biotinylated baits at Arbor Biosciences (Ann Arbor, Michigan, USA), tiling across our exon targets with 2× coverage.

Genomic DNA was extracted using the QIAGEN DNeasy Plant Mini Kit following manufacturer's protocols (QIAGEN, Valencia, CA, USA), or a modified CTAB procedure (Doyle & Doyle, 1987). Fresh samples were sonicated to a target fragment size of 550 bp using a Covaris E220 Focused-ultrasonicator (Wohurn, MA, USA). The remaining samples from herbarium specimens were not sonicated, as they mostly had average fragment sizes < 550 bp. Sequencing libraries were prepared using the Illumina TruSeq Nano HT DNA kit (Illumina Inc., San Diego, CA, USA). DNA libraries were checked for quality using the 2100 Bioanalyzer system (Agilent, Santa Clara, CA, USA), and library concentrations were quantified using a Qubit 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA). Indexed samples were pooled in approximately equal quantities, typically 16 samples per equimolar 500–700-ng pool. Pools were enriched using the custom Cyperaceae-specific baits described above following the manufacturer's protocols for the myBaits kit (v.3). We used hybridization temperatures of 60 or 65 ºC and hybridization times of 16 and 24 h for herbarium and silica-dried fresh material, respectively. Enriched products were PCR amplified for 14 cycles and purified using the QIAquick PCR purification kit (Qiagen). Paired-end libraries were sequenced in four lanes on an Illumina MiSeq (2 × 300 bp; 600 cycle v3) at The Field Museum of Natural History (Chicago, IL, USA). Raw reads are available in the NCBI Sequence Read Archive, BioProject PRJNA553631.

Assembly of nuclear and plastid data

De-multiplexed sequences were quality trimmed (> Q20 in a 5-bp window and default parameters) using Trimmomatic (Bolger *et al.*, 2014). Loci were assembled using the HybPiper pipeline with default parameters and the BWA option (Li & Durbin, 2009) to align reads to target sequences. HybPiper uses SPAdes (Bankevich *et al.*, 2012) to assemble these reads into contigs, which were then aligned to the original target sequences (v.1; Johnson *et al.*, 2016). Summary statistics were obtained using Samtools v.1.8 (Li *et al.*, 2009; Supporting Information, Table S1). When multiple long-length sequences were found during the HybPiper pipeline (see Johnson *et al.*, 2016 for further details) we excluded the whole locus because the sequences may represent paralogous genes, alleles or contaminants. Orthologous sequences were aligned in MAFFT v.7.222 using the function '-auto' (Katoh & Standley, 2013). We obtained three matrices: exons only; introns only and supercontigs (exons + introns).

We removed poorly aligned regions from alignments in Gblocks v.0.91b (Castresana, 2000) using default parameters. All individual exons, introns and supercontig matrices were also concatenated into exon, intron and supercontig supermatrices. Summary statistics for the exon, intron and supercontig matrices were obtained using AMAS (Borowiec, 2016).

We recovered plastid loci by mapping sequence reads to the annotated plastome of *Carex siderosticta* (GenBank accession number NC_027250.1; J. Jung, J. Park and S. Kim, unpublished), extracting the coding sequence (CDS) regions from each gene (Supporting Information, Table S4), if similarity was 95% between the two plastomes, using the mapping and transfer annotations functions in Geneious v.9.1.7 (http://www.geneious.com, Kearse *et al.*, 2012). We used HybPiper with default parameters to extract matrices of exons, introns and supercontigs. We obtained 70 exon, 61 intron and 70 supercontig matrices that were aligned with MAFFT. We removed poorly aligned regions from alignments using Gblocks v.0.91b and default parameters. Matrices were concatenated into a supermatrix (exons: 87 taxa, 55 448 bp; introns: 86 taxa, 177 579 bp; supercontigs: 84 taxa, 252 468 bp). Then, we repeated these analyses including only samples with at least 10 kbp, resulting in matrices of equal length but with 66, 65 and 65 taxa for exons, introns and supercontigs (which combines exon and intron matrices), respectively. Summary statistics for each plastid matrix were obtained using AMAS.

Phylogenetic analyses

Nuclear and plastid phylogenetic trees were inferred under maximum likelihood (ML) after automatic model selection using ModelFinder (Kalyaanamoorthy *et al.*, 2017) in IQtree v.1.4.2 (Nguyen *et al.*, 2015; 1000 ultrafast bootstraps '-bb', '-m TEST'). We also used a method that implements the multispecies coalescent models, in which individual genes are allowed to evolve within the species tree under independent tree topologies to account for gene tree discordance, as implemented in Astral-III v.5.6.1 (Zhang, Sayyari & Mirarab, 2017). We used default analysis parameters to estimate a species tree from individual nuclear gene trees based on the unconcatenated DNA matrices in RAxML v.8.2.9, using the GTRCAT model and 200 fast bootstraps followed by slow ML optimization (default '-fa' search; Stamatakis, 2014).

To evaluate sensitivity of phylogenetic inferences to taxon sampling and alignment quality, we analysed nuclear exon matrices with four different locus or sample exclusions. With the full sampling of 100 individuals, we explored the effect of locus sets that showed percentages of identical sites > 40% and 54.6%, which represented different percentages of missing data. We reduced our sampling to 22 individuals, still representing proportionally the number of species found in each main clade, and we estimated the phylogeny from matrices that had percentages of identical sites > 40% and 70%.

Gene trees and species-tree discordance

We evaluated discordance between exon rooted gene trees for all loci as inferred in the previous steps with RAxML and the concatenated species tree (inferred with IQtree) by computing the level of support and conflict of each shared bipartition following Smith *et al.* (2019), using Matthew G. Johnson's scripts for visualization (https://github.com/mossmatters/MJPythonNotebooks). For conflicting nodes, this approach calculates the proportion of gene trees that show the most common and supported alternative bipartition, all other supported conflicting bipartitions (infrequent topologies) and those that have no support for any conflicting bipartition. Analyses were conducted on exon RAxML trees rooted in ape v.4.1 (Paradis, Claude & Strimmer, 2004) of R v.3.4.0 (R Core Team, 2017).

We also calculated two measures of genealogical concordance, gene concordance factors (gCF) and site concordance factors (sCF), using the '-gcf' and '-scf' options in IQtree v.1.7beta (Nguyen *et al.*, 2015; Minh, Hahn & Lanfear, 2018). Both these approaches characterize incongruence among loci and sites without explicitly modelling the processes underlying incongruence, i.e. without making a claim about lineage sorting, introgression or other processes. We also calculated quartet distance between the concatenated species tree and individual nuclear exon

trees from each gene matrix using =Quartet v.1.0.2 (Sand *et al.*, 2014; Smith, 2019) as a way of quantifying the phylogenetic fidelity of each gene tree.

Nomenclature and taxonomy

We identified the earliest validly published names that would apply to the different possible subgenera implied by our phylogenetic study. We focused primarily on the earliest available subgeneric names in *Carex*, which were published by Rafinesque (1819, 1830, 1840); Dumortier (1827); Heer (1836) and Petermann (1849). We also considered the names coined by Kükenthal, because of their special relevance in the taxonomic history of *Carex*, and the few names coined during the 20th century. Although our phylogenetic tree could be divided into subgenera in a variety of ways, we aimed to propose a subgeneric classification that maintained five of the clades recognized in previous studies and that are here recovered with unprecedented support, while also maintaining morphological coherence of each subgenus as much as possible.

RESULTS

Reads and data

Our Cyperaceae-specific bait kit is effective in capturing the targeted genes in *Carex*, for which it was primarily designed, but also in two distantly related genera of Cyperaceae, *Schoenoplectus* (Rchb.) Palla and *Trichophorum* Pers. In our nuclear dataset, the average number of reads per sample was 399 565 (5892-1 650 759; Supporting Information, Table S1), and the percentage of mapped reads per sample (bait capture efficiency) was 44.8% (4.33–74.94). We recovered all 554 loci (Supporting Information, Table S2), although a few loci derived from the *Cyperus papyrus* transcriptome were not successfully captured for most of the samples. From sequences previously published from an Anchored Hybrid Enrichment study based on an angiosperm-specific bait kit (Léveillé-Bourret *et al.*, 2018a), we were still able to recover most of our nuclear targeted loci (Supporting Information, Table S2). We quantified gene capture success as the percentage of summed captured length of all target loci per individual divided by the summed mean length of all reference loci. However, the average number of captured loci per sample was lower; the percentage of summed captured length of all target loci per individual divided by the summed mean length of all reference loci ranges from 4.5% in *Scirpus pendulus* Muhl. (SRR5314654) to 33.2% in *Khaosokia caricoides* D.A.Simpson, Chayam. & J.Parn. (SRR5314668). In our plastid dataset, the average of mapped reads per sample was 36 017 (0-358 679; Supporting Information, Table S1), which is much higher than in the samples from Léveillé-Bourret *et al.* (2018a).

Summary statistics for the nuclear exon and intron matrices are in Supplementary Table S5. Two hundred and forty-four out of 556 targeted loci were flagged as containing potential paralogue sequences (Supporting Information, Table S6). Most of these flags were found for *C. hamata* Sw. (75), *C. subandrogyna* G.A.Wheeler & Guagl. (65) and *C. baldensis* L. (38). Where more than one sequence was found, we suspected that they might correspond to paralogues (polyploidy in *Carex* is extremely rare). All such matrices were excluded from phylogenetic analyses.

The final edited nuclear dataset contained 100 species (96 *Carex*), with a total of 308 exons and 109 123 potentially parsimony informative sites. On average, 25.51% (0.16–73.01%) of data per individual was missing in exon matrices. For locus sets in which a minimum of 40 and 54.6% of individuals shared all loci, 91 and 29 exons were retained, respectively. In the dataset with a reduced sampling (22 taxa), the locus set in which a minimum of 40 and 70% of individuals shared all loci retained 263 and 118 exon matrices, respectively. Five hundred and forty-three introns were recovered, containing a total of 665 904 potentially parsimony informative sites. On average, 78.84% (42.82–91.78%) of data per individual was missing in intron matrices (Supporting Information, Table S5).

The final plastid dataset consisted of 87 taxa, including 66 *Carex* species, with a total of 70 exons and 5220 potentially parsimony informative sites (Supporting Information, Table S7). On average, the percentage of missing data for the plastid exon matrices was 4.61 (0.00–50.2%) per individual.

Phylogenetic results

Nuclear dataset

Six clades were retrieved using the nuclear dataset of 308 exon matrices (Fig. 2) or the 543 intron matrices (Supporting Information, Fig. S1) for 100 taxa: the Siderostictae, Schoenoxiphium, Unispicate, Uncinia, Vignea and Core Carex clades. Reducing the number of exon loci or the number of terminals reduces support for some of these clades, but most are strongly supported in all analyses (Fig. 3). *Carex* was retrieved as monophyletic with strong support in all analyses [bootstrap support (BS) = 100, local posterior probability (LPP) > 0.90; Fig. 3]. In *Carex*, the Siderostictae clade was retrieved as sister to the rest of *Carex* (BS > 99) in all ML analyses and in the coalescent analysis using the 308 exon matrices. However, LPP support decreases using 91 and 29

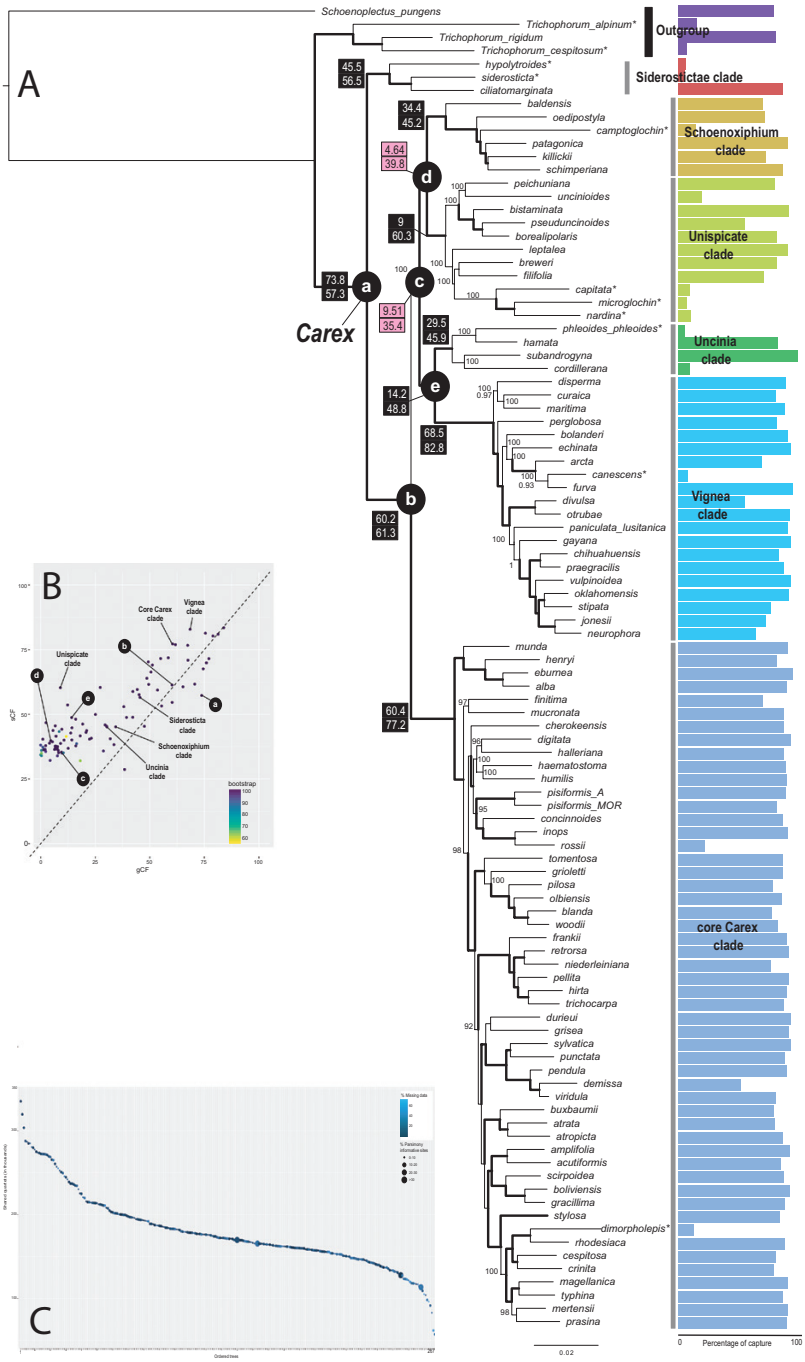

Figure 2. A, Phylogenetic reconstruction obtained under the concatenation approach (maximum likelihood analysis performed with IQtree, 100 samples, 308 concatenated exon loci, 653 347bp), showing the evolutionary relationships within the main clades of *Carex*. Thick branches indicate bootstrap support (BS) = 100 and local posterior probability (LPP) =1, obtained from the coalescent approach in Astral-III; otherwise, branches are not thickened. BS are shown above branches and LPP values below branches; BS and LPP values are only shown when BS > 90 and LPP > 0.9. Numbers on each branch within black boxes show the gene concordance factor (gCF, upper number) and site concordance factor (sCF, lower number); pink boxes highlight the lowest gCF and sCF on nodes C and D. B, Scatter plot of gCF values against sCF values for all branches and they are coloured by BS support. Main nodes are indicated. C, Plot of the individual 267 gene trees ordered upon the number of quartets identically resolved as the reference tree (IQtree). Colours indicate the percentage of missing data, and the size of the point indicates the percentage of potentially parsimony informative sites in each gene tree matrix. Asterisks denote samples from NCBI.

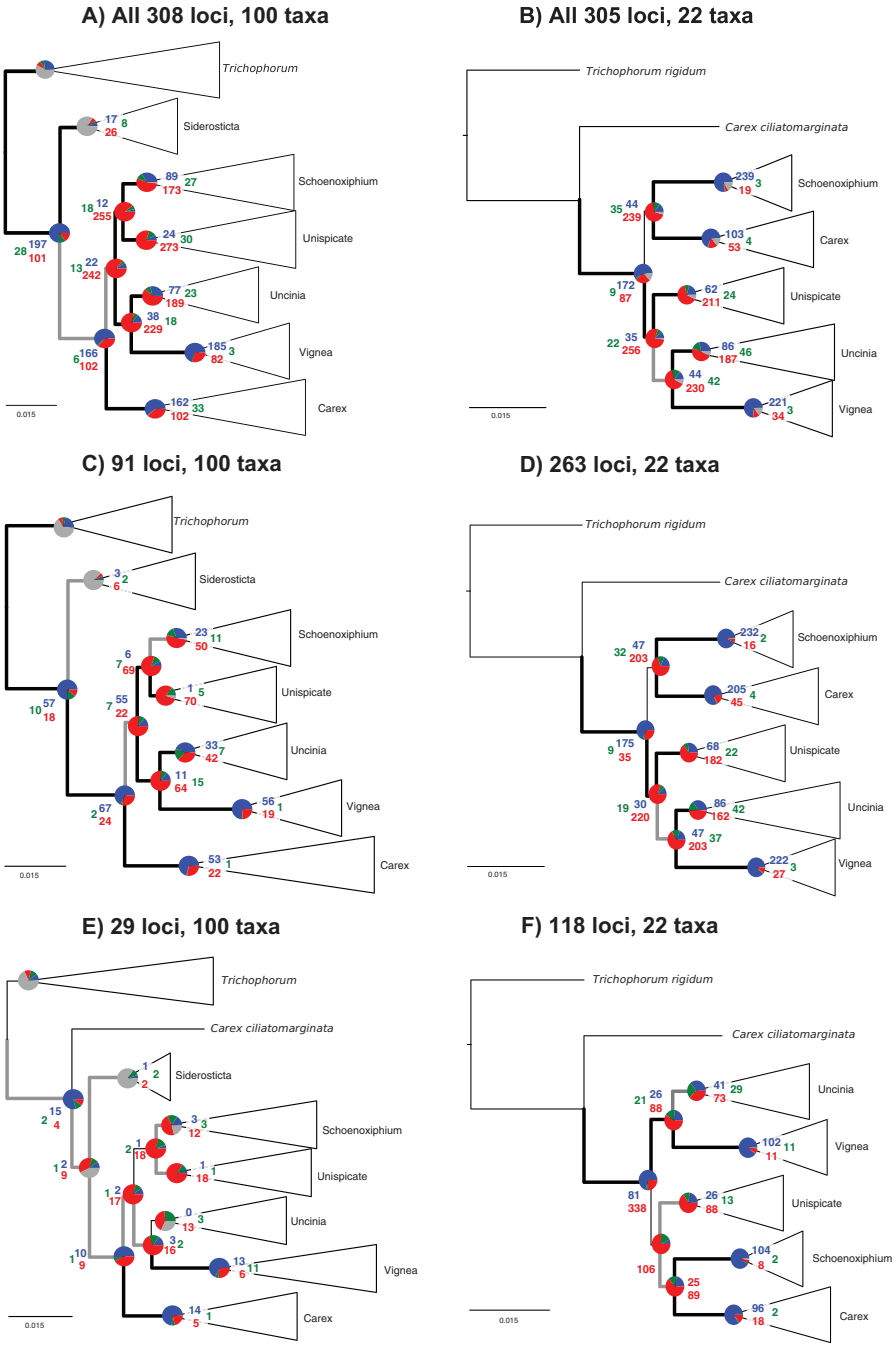

Figure 3. Phylogenetic trees inferred under the concatenation approach (estimated with IQtree) for different number of loci and taxon sampling included in the analyses. A, 308 loci and 100 taxa; B, 305 loci and 100 taxa; C, 91 loci and 100 taxa; D, 263 loci and 100 taxa; E, 29 loci and 100 taxa; F, 118 loci and 100 taxa. Black thickened branches indicate bootstrap support (BS) > 90 and local posterior probability (LPP) > 0.95, obtained from the coalescent approach in Astral-III; grey thickened branches indicate either strong BS (> 90) or LPP (> 0.95) values but not both; unthickened branches indicate lack of support (BS < 0.9 and/or LPP < 0.9). Results from PhyParts of conflicting and concordant gene trees relative to the IQtree are summarized in pie charts in each node. The pie charts represent the proportion of the four categories of concordance and conflict for each node (blue, concordant; green: supporting the dominant alternative topology; red, supporting infrequent remaining alternatives and grey, unsupported). For the main clades in *Carex*, the numbers of gene trees concordant (in blue), supporting an alternative topology (in green) and in conflict (in red), in red with that clade respect the reference tree (IQtree; see Fig. 2).

exon matrices (where a minimum of 40 and 54.6% of individuals shared all loci; Fig. 3C, E). In the reduced dataset of 22 taxa, only one sample of the Siderostictae clade was included (i.e. *C. ciliatomarginata* Nakai) and we could not explore the effect of reducing the exon loci sampling (Fig. 3B, D, F). The Schoenoxiphium and the Unispicate clades appeared as sister clades with strong support (BS = 100, LPP = 1, Fig. 2) with full sampling (100 taxa) and the most complete locus set (308 exon matrices). Support for this clade decreases with the 91 and 29 exon matrices (Fig. 3C, E). The Schoenoxiphium clade was highly supported in the 22-tip datasets (Fig. 3B, D, F). The Unispicate clade is also highly supported when using matrices with 22 taxa, except in the dataset with the minimum number of exon matrices for 22 taxa (118 matrices; Fig. 3F).

Carex cordillerana Saarela & B.A.Ford, *C. hamata, C. phleoides* Cav. and *C. subandrogyna* (the Uncinia clade), which are usually recovered within the Unispicate clade, were resolved as a clade sister to the Vignea clade with very strong support in ML and coalescent analyses (Fig. 2, Supporting Information, Fig. S1). This result has not been previously found in Sanger-sequencing-based studies. Under the full taxon sampling (100 taxa), this relationship is also strongly supported in all analyses (Fig. 3A, C), except for the coalescent analysis using 29 exon matrices, in which it is present but not well-supported (Fig. 3E). In matrices with taxon sampling reduced to 22 taxa, this relationship is highly supported in all ML analyses but not in all coalescent analyses (Fig. 3B, D, F). The relationship between the Schoenoxiphium + Unispicate clade and the Uncinia + Vignea clade was strongly supported in the 100 taxa/308 loci and 100/91 loci analyses and present but not well-supported in the 100/29 loci analysis (Fig. 3A, C, E). *Carex baldensis*, which has usually been found as part of the Unispicate clade (e.g. Global Carex Group, 2016a), was recovered as sister to the Schoenoxiphium clade (Fig. 2, Supporting Information, Fig. S1). The capture success in this species is high and it is retrieved in a strongly supported clade (BS = 100, LPP = 1). The Core Carex clade includes a small early-diverging clade (Fig. 2), with the rest of the Core Carex clade arranged in variously supported subclades roughly mirroring the species arrangement in the Global Carex Group (2016a) tree.

Analyses of matrices trimmed with Gblocks produced similar topologies (results not shown). Supercontig matrices also yielded topologies overall similar to those retrieved with the exon matrices (Supporting Information, Fig. S2). Comparison of the individual exon gene tree topologies to the ML tree (Fig. 3) reveals gene tree concordance for phylogenetic relationships among the main clades. The source of conflict at most nodes is incongruence among low-frequency alternative bipartitions or lack of support for any bipartition, whereas conflict between codominant bipartitions appears to be low (see below).

Gene trees and species-tree discordance

The exon gene trees are concordant with the species tree for most nodes (Fig. 3). Exceptions to this are primarily at nodes dominanted by numerous low-frequency alternative topologies (red portion of the pie charts; Fig. 3). In general, the crown nodes of the main six clades have a greater proportion of exon gene trees concordant with the ML tree (blue portion of the pie charts) than exon gene trees supporting another high-frequency alternative topology (green portion of the pie charts). For the Unispicate clade, this pattern is only found when using reduced sampling (22 taxa), not in the analyses using the full sampling (100 taxa). For the crown node of the Schoenoxiphium + Unispicate clade, which is not recovered using 22 taxa, the number of trees supporting the reference ML tree is always smaller than the number of trees supporting alternative bipartitions. When using 29-exon matrices and 100 taxa, the proportion of trees supporting an alternative topology equals or exceeds the proportion of trees concordant with the reference ML tree for five nodes (i.e. crown nodes of the Siderostictae, Schoenoxiphium, Unispicate and Uncinia clades and the node of the Schoenoxiphium + Unispicate clade; Fig. 3E). In contrast, in the dataset with the fewest individuals sampled but the highest gene coverage (Fig. 3B, D, F), the proportion of gene trees supporting the species-tree topology is always higher than the proportion of gene trees supporting an alternative topology, except for the node grouping the Unispicate, Schoenoxiphium and Core Carex clades, a lineage lacking support in both ML and coalescent analyses (Fig. 3F). The crown node of the Siderostictae clade lacked a strong phylogenetic signal (very large grey portion of the pie charts, Fig. 3A, C, E) primarily due to the low number of captured exon loci for those samples (see Fig. 2).

Gene concordance factor (gCF) and site concordance factor (sCF), performed only with the exon dataset, were highest for the main lineages in *Carex*, except for the Siderostictae and Vignea clades (Fig. 2; Supplementary Table 9). Although all of the backbone nodes have BS values of 100% (Figs 2, 3A), at least two clades display very low values for both concordance factors (<10% for gCF and <33% for sCF; Minh *et al.*, 2018): the crown node of the Schoenoxiphium + Unispicate clade (Node D, gCF = 4.6%, sCF = 39.8%) and the crown node of the Schoenoxiphium + Unispicate + Unicinia + Vignea clade (node C, gCF = 9.51%, sCF = 35.4%; Fig. 2). These two nodes also have the shortest branches in the backbone (Fig. 2, Supporting Information Table S8). Of the 305 and 302 genes analysed for each of

these two branches (nodes C and D, respectively), only 29 and 14 trees (9 and 4.6%), respectively, recover the branch. Fifteen and 16 other single-locus trees support the second best-supported resolution for these nodes (4.92 and 5.3% for nodes C and D, respectively), and eight and three trees support the third best-supported resolution (2.62 and 0.99%, respectively). For node D (crown node of the Schoenoxiphium + Unispicate clade), the number of trees supporting an alternative tree topology is higher than the number of trees concordant with the shown topology. The remaining 253 and 269 loci, respectively, have other infrequent topologies, which suggest that there are many single-locus trees with low information content in the exon dataset. The low gCF and sCF values for these two nodes suggest that the individual gene trees contribute little information about this node. The small number of single-locus trees that contain these groupings is low, and only about a third of the sites are informative for these branches. These low values may indicate, first, that the information gathered in the individual gene trees is insufficient to resolve those branches, or second, that these gene trees show a discordant signal. Such conflicting signal may come from processes such as incomplete lineage sorting or ancient hybridization.

The crown node for the Uncinia clade has a gCF of 29.49% and sCF of 45.92%. From a total of 295 single-locus trees with appropriate taxon sampling to test for the existence of this branch, 132 have all low-frequency topologies and are consequently inconclusive, whereas 87 support the branch. Alternative secondary topologies are strongly supported by 52 and 54 loci, respectively. For the node including the Uncinia + Vignea clade (node E, Fig. 2), 42 trees support the existence of the node; an alternative topology is supported by 21 trees (7.09%). For this node, a total of 233 remaining single-locus trees are inconclusive.

We also quantified the quartet distance from each gene tree to the ML concatenated tree (Fig. 2) to identify loci resolving the highest proportion of bipartitions similarly to the species tree. With only 80 exon loci (Supporting Information, Figs S3, S4), we recovered a topology similar to that obtained with 308 concatenated exon matrices in IQtree.

Plastid dataset

Topologies obtained from nuclear and plastid datasets (exons) using ML mostly resolved similar topologies in each main clade. The most striking difference in placement was the position of the Uncinia clade, which nested in the Unispicate clade with the plastid dataset (Fig. 4) rather than sister to the Vignea clade as found in the nuclear dataset (Figs 2, 3). The hypothesized position of outgroup species was congruent with previous studies (see references in Table 1) and strongly supported here using 66 exon matrices; introns and supercontigs were not necessary for strong resolution (Supporting Information, Figs S5, S6). In our reconstruction, *Carex* formed a well-supported clade (BS = 100; Fig. 4) sister to *Trichophorum*. In *Carex*, the Siderostictae clade (BS = 100) was retrieved as sister to the rest of the genus. The remainder of *Carex* resolved into two main clades, one with the Schoenoxiphium clade (represented only by two accessions) sister to the Unispicate clade (which included the Uncinia clade), the other with the Vignea clade sister to the Core Carex clade. The Core Carex clade was subdivided into a small, strongly supported clade (BS = 100) comparable to the 'Small Core Carex clade' as in Starr *et al.* (2015), and a larger strongly supported clade (BS = 100) comprising all remaining species. The structure of these two internal clades roughly agreed with the Global Carex Group (2016a) topology, which was based in part on plastid (*matK*) data.

Nomenclature and taxonomy

The six major clades recovered in our phylogenetic reconstructions have available names except the Siderostictae clade (see Taxonomic Treatment). The traditional *Carex* subgenus *Carex* (lectotype species: *C. hirta* L.) and subgenus *Vignea* (lectotype species: *C. arenaria* L.) serve as subgeneric names for the Core Carex clade and Vignea clade, respectively. Regarding the other taxa of the Caricoid clade and its possible subdivisions, all available names were published simultaneously, although *C.* subgenus *Psyllophorae* (type species: *C. pulicaris* L.) may be argued to have priority due to its earliest acceptance (Egorova, 1999). In its narrowest conception, subgenus *Psyllophorae* would apply only to the Schoenoxiphium clade. For the Unispicate clade, the name *C.* subgenus *Euthyceras* Peterm. (type species: *C. microglochin* Wahlenb.) has priority, and for the clade grouping the former genus *Uncinia* and its allies, the combination *C.* subgenus *Uncinia* (Pers.) Peterm. (lectotype species: *C. uncinata* L.f.) is available.

The name *C.* subgenus *Caricotypus* Dumort. is to be considered a *nomen invalidum*, since it intentionally includes the type of the genus (McNeill pers. comm.; Turland *et al.*, 2018, ICN, Art. 22.2). Other old forgotten names from Rafinesque (1819, 1830, 1840) and Dumortier (1823) are typified on species in subgenus *Carex* (see Taxonomic Treatment).

DISCUSSION

We present the first phylogenomic analyses of *Carex* based on a nuclear-locus bait kit designed from Cyperaceae transcriptomes to sequence 554 genes.

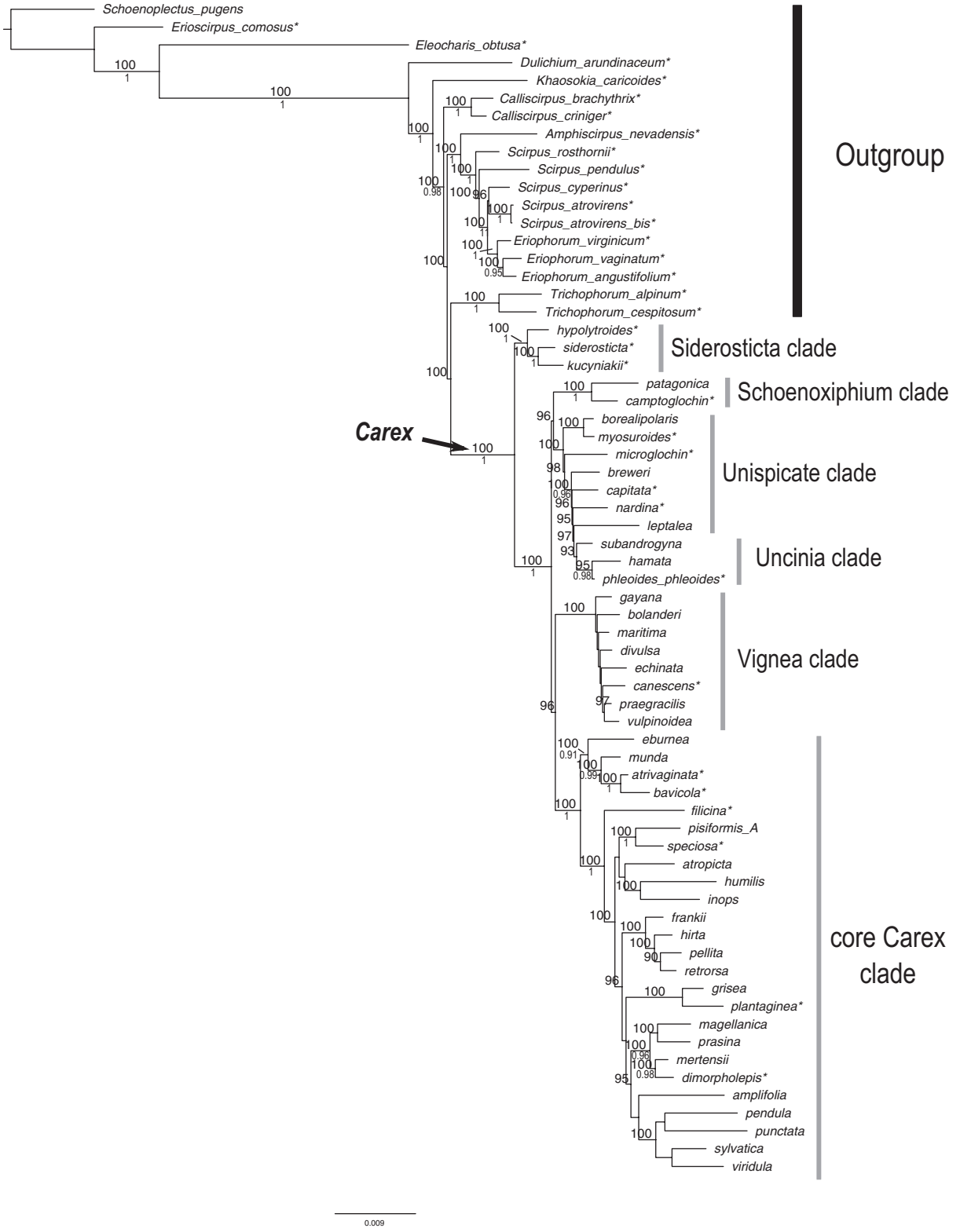

Figure 4. Phylogenetic reconstruction inferred from the plastid dataset (66 species, 70 loci, 55 448bp) under the concatenation approach (maximum likelihood analysis performed with IQtree), showing the evolutionary relationships in *Carex* and closely related genera. Numbers above branches indicate bootstrap (BS > 90) support values. Asterisks denote samples from NCBI.

We target the full phylogenetic breadth of this large (*c*. 2000 spp.) cosmopolitan angiosperm genus with a matrix of 308 nuclear exons, one of 543 nuclear introns and one of 66 plastid exons. Our results demonstrate that there are six strongly supported main lineages in *Carex* (the Siderostictae, Schoenoxiphium, Unispicate, Uncinia, Vignea and Core Carex clades). Support for the hypothesized relationships among these main lineages is the strongest obtained to date.

Previous analyses of relationships among the main *Carex* lineages exhibit a wide range of topological resolutions that are largely incongruent with one another (Table 1; Global Carex Group, 2015). We here demonstrate that an increased number of loci in both nuclear and plastid datasets and the selection of representative samples of the main lineages of *Carex* significantly increase our understanding of relationships among lineages (Figs 2–4). By quantifying the quartet distance from each nuclear exon gene tree to the nuclear ML concatenated tree (Fig. 2), we were able to identify 80 exon loci resolving the highest proportion of bipartitions similar to the species tree. This subset of 'best' nuclear loci may help to identify loci for future barcoding projects in *Carex*.

The narrow time-frame during which the main clades diverged, as inferred by the short branch lengths, resulted in sparse phylogenetic signal for those relationships and may be the most significant factor that has hampered the reconstruction of the backbone of *Carex* in previous studies. Our results show short internal branches subtending the Schoenoxiphium + Unispicate clade (node D, Fig. 2, Supporting Information Fig. S1) and the Schoenoxiphium + Unispicate + Uncinia + Vignea clade (node C, Fig. 2, Supporting Information Fig. S1). These nodes, which are strongly supported in the exon ML analyses but not in the exon coalescent analyses (although strongly supported in the intron analyses), might have suffered from incomplete lineage sorting (i.e. the retention of ancestral polymorphism over successive speciation events; e.g.Degnan & Ronsenberg, 2009) during early diversification of the genus. We also demonstrate that many of the exon matrices have little phylogenetic signal, which may have mundane methodological causes such as high amounts of missing data, or biologically interesting causes such as rapid diversification, extinction or unsampled critical lineages, ancient introgression or incomplete lineage sorting. Thus, the exon gene tree conflict observed here provides a window on the processes that have shaped diversification in *Carex*. A few key phylogenetically isolated Asian taxa with problematic placements (i.e. *C. bostrychostigma* Maxim., *C. dissitiflora* Franch. and *C. satsumensis* Franch. & Sav.) are missing from the present study, and further work is needed to elucidate their placement in one of the retrieved major lineages, or identify whether, to the contrary, they form additional deep branches on the *Carex* tree.

The results of this study represent a signficant step forward in (1) our understanding of Cyperaceae evolution by resolving the relationships of the major lineages in its largest genus, (2) methodology for studying this important plant family through presentation and chararacterization of a novel bait kit for hundreds of genes and (3) a case study in implementing a taxon-specific sequence capture toolkit, demonstrating how robust these kits may be to modifications for efficiency (e.g. reducing the number of genes sequenced) and different analytical pipelines.

A NOVEL CYPERACEAE-SPECIFIC BAIT KIT THAT HELPS RESOLVE RELATIONSHIPS AT SHALLOW EVOLUTIONARY LEVELS

Our Cyperaceae-specific targeted sequencing bait kit successfully resolves species relationships in the extremely diverse genus *Carex*, and it may also help to resolve evolutionary histories in the family at a range of taxonomic levels (e.g. in other genera of Cyperaceae or at tribal or generic level; Larridon *et al.*, 2020). The resolution obtained in *Carex* is higher than any obtained in previous studies (see Table 1) and is already being used to address biogeography, molecular dating and diversification of the genus (Martín-Bravo *et al.*, 2019). Importantly, our Cyperaceae-specific bait kit also shares loci with an angiosperm kit that has become more widely adopted (Buddenhagen *et al.*, 2016; Crowl *et al.*, 2019), enabling us to include and compile data from a previous study (Léveillé-Bourret *et al.*, 2018a). We propose it as the next-generation molecular toolkit for the Cyperaceae community, anticipating that its use will most effectively unite efforts from laboratories worldwide.

In general, the inferred nuclear phylogenetic trees support topologies found in previous studies using Sanger sequencing (Table 1) but with far stronger support, especially for early-branching events. The increase in the number of loci used in the nuclear dataset aided in resolving all the main clades. Our work also revealed previously unforeseen relationships with strong support: a combined Schoenoxiphium + Unispicate clade (node D, Fig. 2, Supporting Information Fig. S1) and a combined Uncinia + Vignea clade (node E, Fig. 2, Supporting Information Fig. S1). All previous phylogenetic reconstructions of *Carex* have recovered the Uncinia clade and *C. baldensis* in the Unispicate clade (e.g. Global Carex Group, 2016a). However, this pattern is unsupported in our nuclear topologies (Fig. 2, Supporting Information Fig. S1), in which the Uncinia clade appears as sister to the Vignea clade and

C. baldensis is recovered as sister to members of the Schoenoxiphium clade.

While the agreement between Sanger-based and genomic plastid phylogenies is not surprising [early studies based on Sanger sequencing represent resolutions of a a single non-recombinant genome (in plastid DNA-only studies) or small number of loci with generally a strong plastid contribution], it is remarkable that the particular topology so robustly found in our study was never previously recovered. One explanation for this phenomenon may be a pervasive but low rate of phylogenetic incongruence in the multiple-copy nuclear ribosomal markers more widely used in phylogenetic analyses of Cyperaceae (i.e. ETS and ITS). At least for ITS, it has been demonstrated that as many as 35 different copies/variants, some of them non-functional variants, in different individuals may be found within a single angiosperm species (King & Roalson, 2008; Song *et al.*, 2012), of which only the predominant copy or copies are likely to be recovered using Sanger sequencing (Nieto-Feliner & Roselló, 2007). It may be possible that through concerted evolution the predominant copies have become those conflicting with the coalescent history of the nuclear genome.

The subgeneric delimitation of *Carex*

Infrageneric classification of *Carex* has been intensely debated since the first molecular phylogenetic study of the genus *c*. 20 years ago (Table 1). While for historical and practical reasons a large part of the discussion has revolved around the sectional arrangement of the genus (Reznicek, 1990; Global Carex Group, 2016a), the nature of the subgenera of *Carex* has drawn the attention of cyperologists, as these large clades sketch out the major macroevolutionary events in the genus. The phylogenetic framework suggested by the inflorescence configuration in *Carex* was recognized very early in morphological treatments, starting with Tuckerman (1843) and Drejer (1844), and culminating in Kükenthal's (1909) monograph, which had a major influence on the classification we use today. Kükenthal's work organized the genus by perigynium structure and inflorescence morphology. Subsequent authors proposed that several of the subgenera might not be monophyletic (Raymond, 1959; Koyama, 1962; Reznicek, 1990; Egorova, 1999). In fact, only subgenus *Vignea* evaded substantial debate because of its remarkable morphological homogeneity.

The first phylogenetic analyses based on DNA sequences demonstrated that classification of the genus might be more complicated than imagined. It was quickly apparent, for example, that the other genera of Cariceae were nested in *Carex* (Starr *et al.*, 1999; Yen & Olmstead, 2000; Roalson *et al.*, 2001). It soon became clear that none of the subgenera proposed to date was entirely monophyletic: *Psyllophorae* was spread among four of the larger clades; *Vigneastra* among five clades; subgenus *Carex* among three clades and *Vignea*, despite its morphological homogeneity, turned out not to include *C. baldensis* and *C. curvula*. In the face of this taxonomic chaos, Molina, Acedo & Llamas (2012) were nonetheless able to trace common patterns and evolutionary trends in inflorescence configuration for all major clades except the Siderostictae clade, providing some of the morphological background necessary for raising of all major clades as formal taxonomic groups.

Our phylogenetic results provide further evidence of the infrageneric arrangement of *Carex*. Whereas all previous studies have pointed to the recognition of four or five major groups (depending on whether the Unispicate and Schoenoxhiphium clades are considered as a single Caricoid clade or separate clades), our work demonstrates the existence of at least six different groups: the five major clades so far retrieved, plus a sixth Uncinia clade. According to the phylogenomic evidence presented here, taken into account with the species groupings resolved by the Global Carex Group (2016a) and Martín-Bravo *et al.* (2019), we propose the following infrageneric classification based on the six main lineages of *Carex*:

1) Core Carex clade: to be treated as the type subgenus, as it includes the lectotype species *C. hirta* L. It also includes the majority of the species formerly treated as subgenus *Vigneastra*, including the lectotype species *C. indica* L. and a few unispicate groups previously placed in subgenus *Psyllophorae*.
2) Unispicate clade: a taxonomically heterogeneous assemblage in the genus. It contains the vast majority of the unispicate *Carex* spp., a selection of North American species formerly included in subgenus *Vigneastra* (Global Carex Group 2016a) and the species of the formerly recognized genus *Kobresia*.
3) Schoenoxiphium clade: perhaps the most taxonomically and morphologically heterogeneous group, including various species formerly placed in subgenera *Vigneastra* (*C. distachya* Desf.), *Vignea* (*C. curvula*, *C. baldensis*), *Psyllophorae* (e.g. *C. pulicaris*) and *Carex* (e.g. *C. phalaroides* Kunth), plus the formerly recognized genus *Schoenoxiphium*.
4) Siderostictae clade: composed of a few early-diverging lineages formerly placed in subgenera *Carex* and *Vigneastra*. It is the only major group of *Carex* with a rather limited geographical area, confined entirely to Eastern Asia.
5) Uncinia clade: comprising the species of former genus *Uncinia* and several closely related lineages formerly included in subgenus *Psyllophorae* and

some Neotropical species formerly included in subgenus *Vigneastra* and the formerly recognized monotypic genus *Cymophyllus*.

6) Vignea clade: largely unchanged after the exclusion of *C. baldensis* and *C. curvula* and the inclusion of a number of dioecious species. According to Yano *et al.* (2014), it should also include *C. satsumensis*, previously included in subgenus *Vigneastra*.

The following taxonomic treatment details the main characters, distribution and composition of each group.

Taxonomic treatment

The following taxonomic proposal is prepared to reflect the evolutionary history of the main *Carex* lineages, rather than to arrange *Carex* diversity into morphologically homogeneous groups. The early use of subgenera and sections was intended to organize the huge diversity displayed by *Carex*. With the integration of an evolutionary perspective into the systematic scheme for *Carex*, starting with Egorova (1999), subgenera in *Carex* were understood as a translation of the evolutionary relationships in the genus to a formal nomenclatural scheme. In this respect, Ball & Reznicek (2002) were pioneers, since they explicitly did not present a subgeneric treatment because of the homoplasy that certain groups presented, although they ordered the sections according to morphological affinities (thus, roughly using the classically recognized subgenera). They foresaw the situation that we are addressing in this work, in which we aim to put to an end to the instability of the subgeneric system that caricologists have been coping with for more than two decades. Given the recurrent homoplasy observed in *Carex* (Global Carex Group, 2016a), making well-defined groups from a phenetic perspective is impractical. Our subgeneric proposal accounts for the six largest clades detected by our phylogenetic results. Despite comprising more or less morphologically similar species, each of the proposed subgenera cannot be readily defined by morphological characters alone: indeed, each subgenus includes a number of taxa that show exceptional morphological characters in relation to their group, reflecting the great morphological diversification of *Carex*.

The following taxonomic synopsis is not intended to be an exhaustive list of subgeneric names published in *Carex*. Rather, we provide the most relevant names in terms of priority and common usage. To avoid the deleterious effects of some of the older forgotten names from Rafinesque (1819, 1830, 1840) and Dumortier (1827) assuming priority over any of the names already in use, we here typify those names on species in the type subgenus *Carex*. By these actions, we intend to help stabilize the infrageneric classification of *Carex*, the nomenclature of which has not undergone major changes since 1999 (see Egorova, 1999). Species included under each of the six subgenera are provided in Supporting Information (Table S10).

Carex L., Sp. Pl. 2: 972. 1753.

Type: *Carex hirta* L. (lectotype designated by Green, 1930).

1. *Carex* subgenus *Siderosticta* M.J.Waterway, subg. nov.

Type: *Carex siderosticta* Hance

Diagnosis: plants woody-rhizomatous, bearing leafy vegetative shoots and pseudolateral (rarely central) reproductive culms with leaves reduced to spathelike bracts. Leaves generally broad, mostly up to 30 mm wide but ranging up to 120 mm wide in some pseudopetiolate species, less commonly as narrow as 2–3 mm. Inflorescence racemose or paniculiform, with two to many spikes; spikes bisexual and androgynous, borne singly or in fascicles or on higher order branches, or rarely unisexual and then with staminate spikes borne distal to the pistillate spikes on each culm, each spike with a more or less developed peduncle subtended by an expanded, spathelike bract, bladeless or with short blade; cladoprophylls on penultimate branches tubular or utriculiform; stigmas three; perigynia closed forming utricles; rachilla absent; chromosome numbers $2n = 12$ or 24.

Species included: 30 species (Supporting Information, Table S10).

Distribution: endemic to eastern Asia, with a centre of diversity in south-eastern Asia.

Notes: The placement of species from section *Hypolytroides* Nelmes (*C. hypolytroides* Ridl. and *C. moupinensis* Franch.) in this subgenus should be considered provisional, since previous reconstructions suggest that section may be sister to the rest of subgenus *Siderosticta*, and it is a morphologically divergent group. Additional data is required to confirm the position of section *Hypolytroides* among members of subgenus *Siderosticta*.

2. *Carex* subgenus Carex

≡ *Carex* subgenus *Dichostachys* Dumort., Fl. Belg. 146. 1827.

Lectotype (here designated): *Carex hirta* L.

= *Carex* subgenus *Scuria* Raf., J. Phys. Chim. Hist. Nat. Arts 89: 106. 1819.

Lectotype (here designated): *Carex lenticularis* Michx.

= *Carex* subgenus *Triodex* Raf., J. Phys. Chim. Hist. Nat. Arts 89: 106. 1819.

Lectotype (here designated): *Carex flava* L.

= *Carex* subgenus *Triplima* Raf., J. Phys. Chim. Hist. Nat. Arts 89: 106. 1819.

Lectotype (here designated): *Carex grisea* Wahlenb.

= *Carex* subgenus *Distinax* (Raf.) Raf., Linnaea 8 (Litt.-Ber.): 83. 1833 = *Carex* section *Distinax* Raf., Bull. Bot. (Geneva) 8: 219. 1830. [basionym].

Lectotype (here designated): *Carex petricosa* Dewey

= *Carex* subgenus *Lentex* (Raf.) Raf., Linnaea 8 (Litt.-Ber.): 83. 1833 = *Carex* section *Lentex* Raf., Bull. Bot. (Geneva) 8: 219. 1830. [basionym].

Lectotype (here designated): *Carex lenticularis* Michx.

= *Carex* subgenus *Onatex* (Raf.) Raf., Linnaea 8 (Litt.-Ber.): 83. 1833 = *Carex* section *Onatex* Raf., Bull. Bot. (Geneva) 8: 219. 1830. [basionym].

Lectotype (here designated): *Carex subspathacea* Wormsk.

= *Carex* subgenus *Tristimex* (Raf.) Raf., Linnaea 8 (Litt.-Ber.): 83. 1833 = *Carex* section *Tristimex* Raf., Bull. Bot. (Geneva) 8: 218. 1830. [basionym].

Lectotype (here designated): *Carex flava* L.

= *Carex* subgenus *Planarex* Raf., Good Book: 25. 1840.

Lectotype *Carex nigra* All. (designated by E.D. Merrill, Index Raf. 79. 1949) (=*Carex parviflora* Host).

= *Carex* subgenus *Planeuris* Raf., Good Book: 25. 1840.

Lectotype *Carex extensa* Good. (designated by E.D. Merrill, Index Raf. 79. 1949).

= *Carex* subgenus *Vigneastra* (Tuck.) Kük., Bot. Jahrb. Syst. 27: 516. 15. 1899 = *Carex* sect. *Vigneastra* Tuck., Enum. Meth. Caric. 10, 18. 1843. [basionym].

Lectotype: *Carex indica* L. (designated by Kern & Nooteboom in C.G.G.J. van Steenis (ed.). Fl. Males. 9: 117. 1979).

= *Carex* subgenus *Indocarex* Baill. ex Kük. in H.G.A. Engler, Pflanzenr. IV, 20(Heft 38): 68, 251. 1909.

Lectotype: *Carex indica* L. (designated by Börner, Abh. Naturwiss. Vereins Bremen 21: 275. 1912).

= *Carex* subg. *Altericarex* H. St. John & C. S. Parker, Am. J. Bot. 12: 66. 1925.

Type: *Carex concinnoides* Mack.

= *Carex* subgenus *Kuekenthalia* Savile & Calder, Can. J. Bot. 31: 171. 1953.

Type: *Carex vesicaria* L.

= *Carex* subgenus *Kreczetoviczia* T.V.Egorova, Bot. Zhurn. (Moscow & Leningrad) 70: 1554. 1985.

Type: *Carex aquatilis* Wahlenb.

Diagnosis: inflorescence racemose or paniculiform, with two to many spikes, rarely reduced to a single spike; spikes unisexual, androgynous or gynaecandrous, with a more or less developed peduncle, rarely sessile; cladoprophylls on the second-to-last branches tubular or utriculiform, rarely reduced or absent (in certain species, some of which with the entire inflorescence reduced to a single spike); stigmas two or three (four in *C. concinnoides*); perigynia closed, forming utricles; rachilla absent.

Species included: 1374 species (Supporting Information, Table S10)

Distribution: cosmopolitan, only absent from Antarctica.

Notes: although the position of *C. bostrychostigma* and *C. dissitiflora* seems uncertain in light of the most recent phylogenetic trees (Global Carex Group, 2016a; Martín-Bravo *et al.*, 2019), their morphological affinities and previous data (Waterway *et al.*, 2015) suggest their provisional placement in subgenus *Carex* until more data are available.

3. ***Carex* subgenus *Euthyceras*** Peterm., Fl. Deutschl. 602. 1849.

Type: *Carex microglochin* Wahlenb. [subgenus described as monotypic]

Diagnosis: Inflorescence reduced to a single spike or more or less paniculiform, rarely racemose and spike-like or forming a congested head (some species in the former genus *Kobresia*); spikes androgynous or gynaecandrous, rarely unisexual (in some species including the species placed in the former genus *Kobresia*), sessile or peduncled; cladoprophylls absent or those on the penultimate branches very differently developed, from scale-like to tubular, rarely funnel-like or more or less utriculiform; stigmas two or three; perigynia closed, forming utricles or open and scale-like, sometimes the margins only fused at base; rachilla absent or present, entirely contained within the perigynium or protruding from its apex, sterile and straight or bearing distally male flowers (former genus *Kobresia*).

Species included: 124 species (Supporting Information, Table S10)

Distribution: widespread in the Northern Hemisphere, with disjunct centres of diversity in North America and the Himalayas, reaching the Southern Hemisphere only in South America and New Zealand.

4. ***Carex* subgenus *Psyllophorae*** (Degland) Peterm., Fl. Deutschl. 602. 1849.

≡ *Psyllophorae* Degland in J.-L.-A. Loiseleur-Deslongchamps, Fl. Gallica, ed. 2, 2: 282. 1828. [basionym].

Type: *Carex psyllophora* L.f. (=*C. pulicaris* Lightf.).

= *Carex* subgenus *Primocarex* Kük. in H.G.A. Engler, Pflanzenr. IV, 20 (Heft 38): 68. 1909, nom. illeg.

Type: *Carex pulicaris* Lightf.

Diagnosis: inflorescence reduced to a single spike or more or less paniculiform, rarely racemose and spike-like or forming a congested head (*C. baldensis*, *C. curvula*); spikes androgynous, rarely unisexual, sessile or peduncled; cladoprophylls absent or those on

the penultimate branches differently developed, from scale-like to tubular, rarely funnel-like or more or less utriculiform; stigmas two or three; perigynia closed forming utricles; rachilla absent or present, entirely contained within the perigynium or protruding from its apex, sterile and straight or bearing distally male flowers (former genus *Schoenoxiphium*).

Species included: 53 species (Supporting Information, Table S10).

Distribution: Western Palearctic and Southern Hemisphere (South America, sub-Saharan Africa, and New Zealand), reaching the Arabian Peninsula, with centres of diversity in Patagonia and Cape Region.

5. ***Carex* subgenus *Uncinia*** (Pers.) Peterm., Fl. Deutschl. 602. 1849.

≡*Uncinia* Pers., Syn. Pl. 2: 534. 1807. [basionym].

Type: *Uncinia australis* Persoon, nom. illeg. (=*Carex uncinata* L.f.) (designated by Pfeiffer, 1874–1875, Nomenclator Botanicus 2 (2): 1529).

Diagnosis: inflorescence reduced to a single spike or racemose and spike-like (some Neotropical species); spikes androgynous, peduncled in racemose inflorescences; cladoprophylls absent or those on the penultimate branches funnel-like or more or less utriculiform (some Neotropical species); stigmas three; perigynia closed, forming utricles; rachilla absent or present, entirely contained within the perigynium or protruding from its apex and hooked at the tip (former genus *Uncinia*).

Species included: 99 species (Supporting Information, Table S10).

Distribution: primarily the Neotropics, Australia and New Zealand, but also lineages endemic to North America, the former genus *Uncinia* also present in Malesia and Pacific and sub-Antarctic archipelagos. It is the only group of *Carex* present in a true Antarctic region [*C. meridensis* (Steyerm.) J.R.Starr, South Georgia archipelago]. Disjunct centres of diversity in South America and New Zealand.

6. ***Carex* subgenus *Vignea*** (P.Beauv. ex T.Lestib.) Heer, Mitth. Geb. Theor. Erdk. 1: 426. 1836.

≡ *Vignea* P.Beauv. ex T.Lestib., Essai Cyp. 22. 1819. [basionym].

Type: *Carex arenaria* L. (designated by T.V. Egorova, Bot. Zhurn. (Moscow & Leningrad) 75: 865. 1990).

= *Carex* subgenus *Caricotypus* Dumort., nom. inval., Fl. Belg. 145. 1827.

= *Carex* subgenus *Cyperoideae* (G.Don) Peterm., Fl. Deutschl. 602. 1849.

Type: *Carex cyperoides* L. (= *C. bohemica* Schreb.).

Diagnosis: inflorescence racemose, spike-like, less often paniculiform, with two to many spikes, rarely reduced to a single spike; spikes androgynous or, less often, gynaecandrous, rarely unisexual (in some species like *C. gayana, C. exilis* and *C. simulata*), sessile; cladoprophylls absent, rarely reduced and scale-like or utriculiform; perigynia closed, forming utricles; stigmas mostly two, rarely three; rachilla absent.

Species included: 330 species (Supporting Information, Table S10).

Distribution: cosmopolitan, with a centre of diversity in North America.

Notes: although the position of *C. satsumensis* seems uncertain in the light of the latest phylogenetic analyses (Global Carex Group, 2016a; Martín-Bravo *et al.*, 2019), the Yano *et al.* (2014) trees suggest its placement provisionally under subgenus *Vignea* until more data are available.

ACKNOWLEDGEMENTS

We would like to thank two anonymous reviewers for their comments on the manuscript. I.L. thanks the B.A. Krukoff Fund for its support; B.A.F thanks the National Geographic Society Research and Exploration for the grant number 9035-11, 9441-14; JRS thanks the National Science and Engineering Research Council of Canada, grant number RGPIN-2018–04115; E.M. thanks the Spanish Government, FPU-Fellowship AP2012-2189 and Juan de la Cierva-Formación FJCI-2017–32314. The authors also thank John McNeill and Rafaël Govaerts for discussion on nomenclature; James L. Reveal for sharing the names list; Matthew G. Johnson and Elliot Gardner for their help during laboratory work and data analyses; Kevin Feldheim for his help conducting the laboratory work in the Pritzker Lab at The Field Museum (Chicago, USA) and J.M.G. Cobos for his help with laboratory work.

FUNDING

This work was supported by the US National Science Foundation (grants NSF-DEB awards 1255901 to ALH and MJW and 1256033 to ER) and the Spanish Ministry of Economy and Competitiveness (grant CGL2016-77401-P to ML and SMB). This work was partially supported by the National Research Foundation of Korea[NRF-2017R1D1A1B03034952] to S. Kim.

J. Park helped to produce the transcriptome of *Carex siderosticta* for probe design. I. Larridon, S. Martín-Bravo, R. F. C. Naczi and A. A. Reznicek contributed together with P.J.-M. to the nomenclatural study of *Carex* subgeneric names.
Other authors (listed in alphabetical order): M. Escudero, B. A. Ford, S. Gebauer, M. H. Hoffmann, D. A. Simpson, J. R. Starr and K. L. Wilson.

Affiliations (listed alphabetically by country):
K.L. Wilson, National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Mrs Macquaries Road, Sydney NSW 2000, Australia
I. Larridon, Royal Botanic Gardens, Kew, Richmond, Surry, TW9 3AE, UK and Systematic and Evolutionary Botany Lab, Department of Biology, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium.
S. Gebauer, Martin Luther University Halle-Wittenberg, Neuwerk 21, 06108 Halle, Germany.
M. H. Hoffmann, Martin Luther University Halle-Wittenberg, Neuwerk 21, 06108 Halle, Germany.
D. A. Simpson, Royal Botanic Gardens, Kew, Richmond Surrey TW9 3AE, UK; and Department of Botany, School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland.
R. F. C. Naczi, New York Boanical Garden, 2900 Southern Blvd., Bronx, NY 10458–5126, USA.
A. A. Reznicek, University of Michigan Herbarium, 3600 Varsity Drive, Ann Arbor, MI 48108-2228, USA.
B. A. Ford, Department of Biological Sciences, University of Manitoba, Winnipeg, R3T 2N2, Manitoba, Canada.
J. R. Starr, Department of Biology, University of Ottawa, Ottawa, Ontario, K1N 6N5, Canada.
J. Park, InfoBoss Co. Ltd, Seoul 06088, Korea.
M. Escudero, Department of Plant Biology and Ecology, University of Seville, Reina Mercedes sn, ES-41012 Seville, Spain.
S. Martín-Bravo, Department of Molecular Biology and Biochemical Engineering, Universidad Pablo de Olavide, ctra de Utrera km 1 sn, 41013 Seville, Spain.

REFERENCES

Allan GJ, **Porter JM. 2000.** Tribal delimitation and phylogenetic relationships of Loteae and Coronilleae (Faboideae: Fabaceae) with special reference to *Lotus*: evidence from nuclear ribosomal ITS sequences. *American Journal of Botany* **87:** 1871–1881.

APG IV. **2016.** An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181:** 1–20.

Bakker FT. 2017. Herbarium genomics: skimming and plastomics from archival specimens. *Webbia*, **72:** 35–45.

Ball PW, **Reznicek AA. 2002.** *Carex*. In: Flora of North America Editorial Committee, ed. *Flora of North America North of Mexico, Vol. 23. Magnoliophyta: Commelinidae (in part): Cyperaceae*. New York: Oxford University Press, 254–272.

Banasiak Ł, **Wojewódzka A**, **Baczyński J**, **Reduron JP**, **Piwczyński M**, **Kurzyna–Młynik R**, **Gutaker, R**, **Czarnocka-Cieciura A**, **Kosmala-Grzechnik S**, **Spalik K. 2016.** Phylogeny of Apiaceae subtribe Daucinae and the taxonomic delineation of its genera. *Taxon* **65:** 563–585.

Bankevich A, **Nurk S**, **Antipov D**, **Gurevich AA**, **Dvorkin M**, **Kulikov AS**, **Lesin VM**, **Nikolenko SI**, **Pham S**, **Prjibelski AD**, **Pyshkin AV**, **Sirotkin AV**, **Vyahhi N**, **Tesler G**, **Alekseyev MA**, **Pevzner PA. 2012.** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19:** 455–477.

Bolger AM, **Lohse M**, **Usadel B. 2014.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30:** 2114–2120.

Borowiec ML. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* **4:** e1660.

Buddenhagen C, **Lemmon AR**, **Lemmon EM**, **Bruhl J**, **Cappa J**, **Clement WL**, **Donoghue M**, **Edwards EJ**, **Hipp AL**, **Kortyna M**, **Mitchell N**, **Moore A**, **Prychid CJ**, **Segovia-Salcedo MC**, **Simmons MP**, **Soltis P**, **Wanke S**, **Mast A. 2016.** Anchored phylogenomics of angiosperms I: assessing the robustness of phylogenetic estimates. *bioRxiv* 086298.

Cardoso D, **de Queiroz LP**, **Pennington RT**, **de Lima HC**, **Fonty É**, **Wojciechowski MF**, **Lavin M. 2012.** Revisiting the phylogeny of papilionoid legumes: new insights from comprehensively sampled early-branching lineages. *American Journal of Botany* **99:** 1991–2013.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17:** 540–552.

Chamala S, **García N**, **Godden GT**, **Krishnakumar V**, **Jordon-Thaden IE**, **Smet RD**, **Barbazuk WB**, **Soltis DE**, **Soltis PS. 2015.** MarkerMiner 1.0: a new application for phylogenetic marker development using angiosperm transcriptomes. *Applications in Plant Sciences* **3:** 1400115.

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

Crowl AA, **Manos PS**, **McVay JD**, **Lemmon AR**, **Lemmon EM**, **Hipp AL. 2019.** Uncovering the genomic signature of ancient introgression between white oak lineages (*Quercus*). *New Phytologist* **226:** 1158–1170.

Dai LK, **Liang SY**, **Zhang SR**, **Tang Y**, **Koyama T**, **Tucker GC**, **Simpson DA**, **Noltie HJ**, **Strong MT**, **Bruhl JJ**, **Wilson KL**, **Muasya AM. 2010.** Cyperaceae. In: Wu ZY, Raven PH, Hong DY, eds. *Flora of China*, **Vol. 23**. Beijing, St. Louis: Science Press and Missouri Botanical Garden Press, 164–461.

Degnan JH, **Rosenberg NA. 2009.** Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* **24:** 332–340.

Dorsey BL, **Haevermans T**, **Aubriot X**, **Morawetz JJ**, **Riina R**, **Steinmann VW**, **Berry PE. 2013.** Phylogenetics, morphological evolution, and classification of *Euphorbia* subgenus *Euphorbia*. *Taxon* **62:** 291–315.

Downie SR, **Spalik K**, **Katz–Downie DS**, **Reduron JP. 2010.** Major clades within Apiaceae subfamily Apioideae as inferred by phylogenetic analysis of nrDNA ITS sequences. *Plant Diversity and Evolution* **128:** 111–136.

Doyle JJ, **Doyle JL. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, Botanical Society of America* **19:** 11–15.

Drejer S. 1844. *Symbolae Caricologicae ad synonymiam Caricum extricandam stabiliendamque et affinitates naturales eruendas.* Copenhagen: Bianco Luno.

Dumortier BCJ. 1827. *Florula Belgica.* Tournai: J. Casterman.

Egorova TV. 1999. *The sedges (*Carex *L.) of Russia and adjacent states.* St. Petersburg, St. Louis: St. Petersburg State Chemical–Pharmaceutical Academy and Missouri Botanical Garden Press.

de la Estrella M, **Forest F**, **Klitgård B**, **Lewis GP**, **Mackinder BA**, **de Queiroz LP**, **Wieringa JJ**, **Bruneau A. 2018.** A new phylogeny-based tribal classification of subfamily Detarioideae, an early branching clade of florally diverse tropical arborescent legumes. *Scientific Reports* **8:** 6884.

De Smet R, **Adams KL**, **Vandepoele K**, **Van Montagu MCE**, **Maere S**, **Van de Peer Y. 2013.** Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants. *Proceedings of the National Academy of Sciences, USA* **110:** 2898–2903.

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

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 ITS, ETS, and *matK* sequences: 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 NCBI GenBank Data. *Systematic Botany* **41:** 529–539.

Hardway TM, **Spalik K**, **Watson MF**, **Katz–Downie DS**, **Downie SR.**, **van Wyk BE**, **Tilney PM. 2004.** Circumscription of Apiaceae tribe Oenantheae. *South African Journal of Botany* **70:** 393–406.

Harrison N, **Kidner CA. 2011.** Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? *Taxon* **60:** 1552–1566.

Heer O. 1836. Die Vegetationsverhältnisse des südöstlichen Theils des Cantons Glarus. *Mittheilungen aus dem Gebiete der Theoretischen Erdkunde* **1:** 279–468.

Hendrichs M, **Michalski S**, **Begerow D**, **Oberwinkler F**, **Hellwig FH. 2004b.** 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. 2004a.** *Carex*, subgenus *Carex* (Cyperaceae)–a phylogenetic approach using ITS sequences. *Plant Systematics and Evolution* **246:** 89–107.

Huerta-Cepas J, **Serra F**, **Bork P. 2016.** ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Molecular Biology and Evolution* **33:** 1635–1638.

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

Jiménez–Mejias P, **Vargas P. 2015.** Taxonomy of the tribe Apieae (Apiaceae) revisited as revealed by molecular phylogenies and morphological characters. *Phytotaxa* **212:** 57–79.

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

Johnson MG, **Pokorny L**, **Dodsworth S**, **Botigué LR**, **Cowan RS**, **Devault A**, **Eiserhardt WL**, **Epitawalage N**, **Forest F**, **Kim JT**, **Leebens-Mack JH**, **Leitch IJ**, **Maurin O**, **Soltis DE**, **Soltis PS**, **Wong GK**, **Baker WJ**, **Wickett N. 2018.** A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* **68:** 594–606.

Kadlec M, **Bellstedt DU**, **Maitre NCL**, **Pirie MD. 2017.** Targeted NGS for species level phylogenomics: 'made to measure' or 'one size fits all'? *PeerJ* **5:** e3569.

Kalyaanamoorthy S, **Minh BQ**, **Wong TK**, **von Haeseler A**, **Jermiin LS. 2017.** ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14:** 587.

Käss E, **Wink M. 1996.** Molecular evolution of the Leguminosae: phylogeny of the three subfamilies based on *rbcL*–sequences. *Biochemical Systematics and Ecology* **24:** 365–378.

Kates HR, **Johnson MG**, **Gardner EM**, **Zerega NJC**, **Wickett NJ. 2018.** Allele phasing has minimal impact on phylogenetic reconstruction from targeted nuclear gene sequences in a case study of *Artocarpus*. *American Journal of Botany* **105:** 404–416.

Katoh K, **Standley DM. 2013.** MAFFT: multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30:** 772–780.

Kearse M, **Moir R**, **Wilson A**, **Stones-Havas S**, **Cheung M**, **Sturrock S**, **Buxton S**, **Cooper A**, **Markowitz S**, **Duran C**, **Thierer T**, **Ashton B**, **Meintjes P**, **Drummond A. 2012.** Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647–1649.

Kilian N, Hand R, von Raab–Straube E (general eds) 2009+ (continuously updated). Cichorieae Systematics Portal. – Published at http://cichorieae.e–taxonomy.net/portal/; accessed [date].

King, MG, **Roalson EH. 2008.** Exploring evolutionary dynamics of nrDNA in *Carex* subgenus *Vignea* (Cyperaceae). *Systematic Botany* **33:** 514–524.

Koyama T. 1962. Classification of the family Cyperaceae (2). *Journal of the Faculty of Science, Tokyo University. Section 3: Botany* **8:** 149–278.

Kriebel R, **Drew BT**, **Drummond CP**, **González-Gallegos JG**, **Celep F**, **Mahdjoub MM**, **Rose JP**, **Xiang C–L**, **Hu G–X**, **Walker JB**, **Lemmon EM**, **Lemmon AR**, **Systma KJ. 2019.** Tracking temporal shifts in area, biomes, and pollinators in

the radiation of *Salvia* (sages) across continents: leveraging anchored hybrid enrichment and targeted sequence data. *American Journal of Botany* **106:** 573–597.

Kükenthal G. 1909. *Cyperaceae–Caricoideae.* Leipzig: Engelmann.

Kukkonen I. 1996. Notes on the treatment of Cyperaceae for *Flora Iranica*. *Annalen des Naturhistorischen Museums in Wien. Serie B für Botanik und Zoologie* 87–95.

Larridon I, **Villaverde T**, **Zuntini AR**, **Pokorny L**, **Brewer GE**, **Epitawalage N**, **Fairlie I**, **Hahn M**, **Kim J**, **Maguilla E**, **Maurin O**, **Xanthos M**, **Hipp AL**, **Forest F**, **Baker WJ. 2020.** Tackling rapid radiations with targeted sequencing. *Frontiers in Plant Science* **10:** 1655.

Lemmon AR, **Emme SA**, **Lemmon EM. 2012.** Anchored hybrid enrichment for massively high–throughput phylogenomics. *Systematic Biology* **61:** 727–744.

Lemmon EM, **Lemmon AR. 2013.** High–throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics* **44:** 99–121.

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**, **Lemmon EM**, **Lemmon AR. 2018a.** Resolving rapid radiations within angiosperm families using anchored phylogenomics. *Systematic Biology* **67:** 94–112.

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

Li H, **Handsaker B**, **Wysoker A**, **Fennell T**, **Ruan J**, **Homer N**, **Marth G**, **Abecasis G**, **Durbin R**, **1000 Genome Project Data Processing Subgroup**. **2009.** The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* **25:** 2078–2079.

LPWG. **2017.** A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *Taxon* **66:** 44–77.

Luceño M, **Escudero M**, **Jiménez-Mejías P. 2008.** Carex L. In: Castroviejo S, Luceño M, Galán A, Jiménez-Mejías, Cabezas F, Medina L, eds. *Flora Iberica Vol. 18. Cyperaceae–Pontederiaceae.* Madrid: Real Jardín Botánico, CSIC, 109–250

Martín-Bravo S, **Jiménez-Mejías P**, **Villaverde T**, **Escudero E**, **Hahn M**, **Spalink D**, **Roalson EH**, **Hipp AL**, **the Global Carex Group**. **2019.** A tale of worldwide success: behind the scenes of *Carex* (Cyperaceae) biogeography and diversification. *Journal of Systematics and Evolution*. 57: 695–718.

Merrill ED. 1949. *Index Rafinesquianum*. Harvard: Arnold Arboretum, Harvard University.

Minh BQ, **Hahn MW**, **Lanfear R. 2018.** New methods to calculate concordance factors for phylogenomic datasets. *bioRxiv*: 487801.

Molina A, **Acedo C**, **Llamas F. 2012.** A comparative study of the inflorescence in the genus *Carex* (Cyperaceae). *Systematic Botany* **37:** 365–381.

Nguyen L-T, **Schmidt HA**, **von Haeseler A**, **Minh BQ. 2015.** IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* **32:** 268–274.

Nicolas AN, **Plunkett GM. 2009.** The demise of subfamily Hydrocotyloideae (Apiaceae) and the re-alignment of its genera across the entire order Apiales. *Molecular Phylogenetics and Evolution* **53:** 134–151.

Nieto-Feliner G, **Rosselló JA. 2007.** Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* **44:** 911–919

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

Peirson JA, **Bruyns PV**, **Riina R**, **Morawetz JJ**, **Berry PE. 2013.** A molecular phylogeny and classification of the largely succulent and mainly African *Euphorbia* subgenus *Athymalus* (Euphorbiaceae). *Taxon* **62:** 1178–1199.

Petermann WL. 1849. *Deutschlands Flora.* Leipzig: Georg Wigand's Verlag.

R Core Team. **2017.** *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Rafinesque CS. 1819. Prodrome des nouveaux genres de plantes observés en 1817 et 1818 dans l'intérieur des États-Unis d'Amérique. *Journal de Physique, de Chimie et d'Historie Naturelle* **89:** 96–107.

Rafinesque CS. 1830. Extrait d'une lettre de M.C.S. Rafinesque, professeys à Philadelphie, adressée à M. le professeur De Candolle (Mai 1830). *Bulletin Botanique (Geneva)* **8:** 214–221.

Rafinesque CS. 1833. Extrait d'une lettre de M. C. S. Rafinesque professeur a Philadelphie, addressee à Mr. le prof. De Candolle (Mai 1830). *Linnaea* **8** (Litteratur**):** 79–86.

Rafinesque CS. 1840. *The good book*. Philadelphia: Eleutherium of Knowledge.

Raymond M. 1959. Carices Indochinenses necnon Siamenses. *Memoires de Jardin Botanique de Montreal* **53:** 1–125.

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

Riina R, **Peirson JA**, **Geltman DV**, **Molero J**, **Frajman B**, **Pahlevani A**, **Barres L**, **Morawetz JJ**, **Salmaki Y**, **Zarre S**, **Kryukov A**, **Bruyns PV**, **Berry PE. 2013.** A worldwide molecular phylogeny and classification of the leafy spurges, *Euphorbia* subgenus *Esula* (Euphorbiaceae). *Taxon* **62:** 316–342.

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.

Sand A, **Holt MK**, **Johansen J**, **Brodal GS**, **Mailund T**, **Pedersen CNS. 2014.** tqDist: a library for computing the quartet and triplet distances between binary or general trees. *Bioinformatics* **30:** 2079–2080.

Sauquet H, **Magallón S. 2018.** Key questions and challenges in angiosperm macroevolution. *New Phytologist* **219:** 1170–1187

Smith MR. 2019. Quartet: comparison of phylogenetic trees using quartet and bipartition measures. doi:10.5281/zenodo.2536318. Geneva: Zenodo.

Song J, **Shi L**, **Li D**, **Sun Y**, **Niu Y**, **Chen Z**, **Luo H**, **Pang X**, **Sun Z**, **Liu C**, **Lv A**, **Deng Y**, **Larson-Rabin Z**, **Wilkinson M**, **Chen S. 2012.** Extensive pyrosequencing reveals frequent intra-genomic variations of internal transcribed spacer regions of nuclear ribosomal DNA. *PLoS One* **7:** e43971.

Soreng RJ, **Peterson PM**, **Romaschenko K**, **Davidse G**, **Zuloaga FO**, **Judziewicz EJ**, **Filgueiras TS**, **Davis JI**, **Morrone O. 2017.** A worldwide phylogenetic classification of the Poaceae (Gramineae). *Journal of Systematics and Evolution* **53:** 117–137.

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

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and postanalysis 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 directions. *The 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: Ford BA, Naczi RCF, eds. *Sedges: uses, diversity, and systematics of the Cyperaceae*. St. Louis: Missouri Botanical Garden Press, 245–265.

Starr JR, **Jansen 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.

Tuckerman E. 1843. *Enumeratio methodica caricum quarundam.* Schenectady: Isaacus Riggs:.

Turland NJ, Wiersema JH, Barrie FR, Greuter W, Hawksworth DL, Herendeen PS, Knapp S, Kusber W-H, Li D-Z, Marhold K, May TW, McNeill J, Monro AM, Prado J, Price MJ, Smith GF, eds. 2018. *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017. Regnum Vegetabile 159*. Glashütten: Koeltz Botanical Books.

Vargas OM, **Ortiz EM**, **Simpson BB. 2017.** Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytologist* **214:** 1736–1750.

Vargas P**, Zardoya R, eds. 2014.** *The tree of life.* Sunderland: Sinauer Associates.

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

Villaverde T, **Pokorny L**, **Olsson S**, **Rincón-Barrado M**, **Johnson MG**, **Gardner EM**, **Wickett NJ**, **Molero J**, **Riina R**, **Sanmartín I. 2018.** Bridging the micro- and macroevolutionary levels in phylogenomics: HybSeq solves relationships from populations to species and above. *New Phytologist* **220:** 636–650.

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

Waterway MJ, **Prado A**, **Bruhl J**, **Gehrke B**, **Hoshino T**, **Jin X-F**, **Luceño M**, **Masaki T**, **Phulphong K**, **Simpson D**, **Wilson K**, **Zhang SR. 2015.** Testing sectional monophyly in *Carex* (Cyperaceae). Botany 2015, Edmonton, AB, Canada, July 25–29.

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. Published on the Internet; http://wcsp.science.kew.org/ Retrieved 01 July 2019.

Weitemier K, **Straub SCK**, **Cronn RC**, **Fishbein M**, **Schmickl R**, **McDonnell A**, **Liston A. 2014.** HybSeq: combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences* **2:** 1400042.

Xie Y, **Wu G**, **Tang J**, **Luo R**, **Patterson J**, **Liu S**, **Huang W**, **He G**, **Gu S**, **Li S**, **Zhou X**, **Lam TW**, **Li Y**, **Xu X**, **Wong GK**, **Wang J. 2014.** SOAPdenovo-Trans: *de novo* transcriptome assembly with short RNA-Seq reads. *Bioinformatics* **30:** 1660–1666.

Yang Y, **Riina R**, **Morawetz JJ**, **Haevermans T**, **Aubriot X**, **Berry PE. 2012.** Molecular phylogenetics and classification of *Euphorbia* subgenus *Chamaesyce* (Euphorbiaceae). *Taxon* **61:** 764–789.

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

Yen AC, **Olmstead RG. 2000.** Molecular systematics of Cyperaceae tribe Cariceae based on two chloroplast DNA regions: *ndhF* and *trnL* intron-intergenic spacer. *Systematic Botany* **25:** 479–494.

Zhang C, **Sayyari E**, **Mirarab S. 2017.** ASTRAL-III: increased scalability and impacts of contracting low support branches. In: RECOMB International Workshop on Comparative Genomics (pp. 53–75). Cham: Springer.

SUPPORTING INFORMATION

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

Table S1. Summary statistics obtained from the sequencing and nuclear and plastid mapping for the 90 samples processed *de novo* for this publication and ten other samples obtained from NCBI. Taxonomic placement (columns A–E) follows Global Carex Group (2016a) for sectional assignment and major retrieved clades, and Govaerts *et al.* (2019+) for accepted names and authorities. Collector and laboratory numbers (columns F–G) are provided when available. The information provided in the rest of the columns correspond to the summary statistics regaring the number of reads used and their mapping per sample.
Table S2. Capture success for each exon and sample in the nuclear and plastid dataset.
Table S3. List of 431 LCNGs targeted genes with information of presence of multiple copies in other angiosperms (De Smet *et al*., 2013) and their function in *Arabidopsis thaliana* (Source: MarkerMiner, Chamala *et al.*, 2015).
Table S4. Coding sequence regions of *C. siderostita* used as targets in HybPiper for the plastid analyses.
Table S5. Summary statistics of the different nuclear exon and intron matrices analysed in the present study. Most data have been retrieved using AMAS (Borowiec, 2016). Normalized Robinson Foulds distance (nRF; Smith *et al.*, 2019) and number of quartets resolved as the reference tree (Huerta-Cepas *et al.*, 2016) are also included.
Table S6. Number of sequences recovered per locus. Numbers indicate no sequences were recovered for that locus (0); one single sequence have been recovered for that particular locus (1); more than one sequence have been recovered for that particular locus (> 1).
Table S7. Summary statistics of the different plastid exon matrices analysed in the present study.
Table S8. Discordance factors results obtained from the method implemented in IQtree to calculate concordance factor for phylogenomic datasets (Minh *et al.*, 2018).
Table S9. List of the 80 loci that resolved the hightest number of quartets as the reference tree.
Table S10. List of accepted species of *Carex* and its subgeneric placement according to WCSP (2019) with minor additions and modifications according to new names and synonyms published at the end of 2019. Those species sequenced as recorded in Martín-Bravo *et al.* (2019) are specifically indicated.
Figure S1. Phylogenetic tree inferred under coalescence approach (estimated with ASTRAL-III) for the 543 nuclear intron matrices. Branch labels indicate local posterior probabilities.
Figure S2. Phylogenetic tree inferred under coalescence approach (estimated with ASTRAL-III) for the 447 nuclear intron matrices. Branch labels indicate local posterior probabilities.
Figure S3. Phylogenetic tree inferred under concatenation approach (estimated with IQtree) for the 80 nuclear exon matrices. Branch labels indicate bootstrap values.
Figure S4. Phylogenetic tree inferred under coalescence approach (estimated with ASTRAL-III) for the 80 nuclear exon matrices. Branch labels indicate local posterior probabilities.
Figure S5. Phylogenetic tree inferred under concatenation approach (estimated with IQtree) for the 65 plastid intron matrices (177 579bp). Branch labels indicate bootstrap values.
Figure S6. Phylogenetic tree inferred under concatenation approach (estimated with IQtree) for the 65 plastid supercontig matrices (252 468bp). Branch labels indicate bootstrap values.