Universitat Autònoma de Barcelona

Dipòsit digital de documents de la UAB

This is the **accepted version** of the journal article:

Herrando-Moraira, Sonia; Calleja, Juan Antonio; Galbany-Casals, Mercè; [et al.]. «Nuclear and plastid DNA phylogeny of tribe Cardueae (Compositae) with Hyb-Seq data : A new subtribal classification and a temporal diversification framework». Molecular Phylogenetics and Evolution, Vol. 137 (August 2019), p. 313-332. DOI 10.1016/j.ympev.2019.05.001

This version is available at https://ddd.uab.cat/record/288609

under the terms of the BY-NC-ND license

Manuscript Details

Manuscript number	MPE_2019_108_R1
Title	Nuclear and plastid DNA phylogeny of tribe Cardueae (Compositae) with Hyb-Seq data: A new subtribal classification and a temporal diversification framework
Article type	Research Paper

Abstract

Classification of tribe Cardueae in natural subtribes has always been a challenge due to the lack of support of some critical branches in previous phylogenies based on traditional Sanger markers. With the aim to propose a new subtribal delimitation, we applied a Hyb-Seq approach to a set of 76 Cardueae species representing all subtribes and informal groups defined in the tribe, targeting 1061 nuclear conserved orthology loci (COS) designed for Compositae and obtaining chloroplast coding regions as by-product of off-target reads. For the extraction of the target nuclear data, we used two strategies, PHYLUCE and HybPiper, and 776 and 1055 COS loci were recovered with each of them, respectively. Additionally, 87 chloroplast genes were assembled and annotated. With three datasets, phylogenetic relationships were reconstructed using both concatenation and coalescent approaches. Phylogenetic analyses of the nuclear datasets fully resolved virtually all nodes with very high support. Nuclear and plastid tree topologies are mostly congruent with a very limited number of incongruent nodes. Based on the well-solved phylogenies obtained, we propose a new taxonomic scheme of 12 monophyletic and morphologically consistent subtribes: Carlininae, Cardopatiinae, Echinopsinae, Dipterocominae (new), Xerantheminae (new), Berardiinae (new), Staehelininae (new), Onopordinae (new), Carduinae (redelimited), Arctiinae (new), Saussureinae (new), and Centaureinae. In addition, we further updated the temporal framework for origin and diversification of these subtribes. Our results highlight the power of Hyb-Seq over Sanger sequencing of a few DNA markers in solving phylogenetic relationships of traditionally difficult groups.

Keywords	Asteraceae; COS targets; Subtribes; Systematics; Phylogenomics; Target enrichment
Taxonomy	Taxonomy, Phylogenetics, Molecular Phylogenetics, Systematics, Botany, Plant Systematics

Corresponding Author	Sonia Herrando-Moraira
Corresponding Author's Institution	Consejo Superior de Investigaciones Cientificas
Order of Authors	Sonia Herrando-Moraira, Juan Calleja, Mercè Galbany-Casals, Núria Garcia-Jacas, Jianquan Liu, Javier López-Alvarado, Jordi López-Pujol, Jennifer Mandel, Sergi Massó, noemi montes, Cristina Roquet, Llorenç Sáez, Alexander Sennikov, Alfonso Susanna, Vilatersana Roser
Suggested reviewers	Vicki Ann Funk, Santiago Ortiz, Christoph Oberprieler, Katy E Jones, Norbert Kilian
Opposed reviewers	Eckhard Von Raab-Straube, Werner Greuter, Frank H Hellwig

Submission Files Included in this PDF

File Name [File Type]

Cover_letter.docx [Cover Letter]

Review_report.docx [Response to Reviewers]

Highlights.docx [Highlights]

Graphical_abstract.pdf [Graphical Abstract]

Main_text_reviewed.docx [Manuscript File]

Figure_1.pdf [Figure]

Figure_2.pdf [Figure]

Figure_3.pdf [Figure]

Figure_4.pdf [Figure]

Figure_5.pdf [Figure]

Figure_6.pdf [Figure]

Figure_7.pdf [Figure]

Figure_8.pdf [Figure]

Supplementary_material.DOCX [Supplementary Material]

Submission Files Not Included in this PDF

File Name [File Type]

APPENDIX_A_calibration_points_random_distributions.R [Supplementary Material]

APPENDIX_B_Selection_calibration_points_from_random_distributions.xlsx [Supplementary Material]

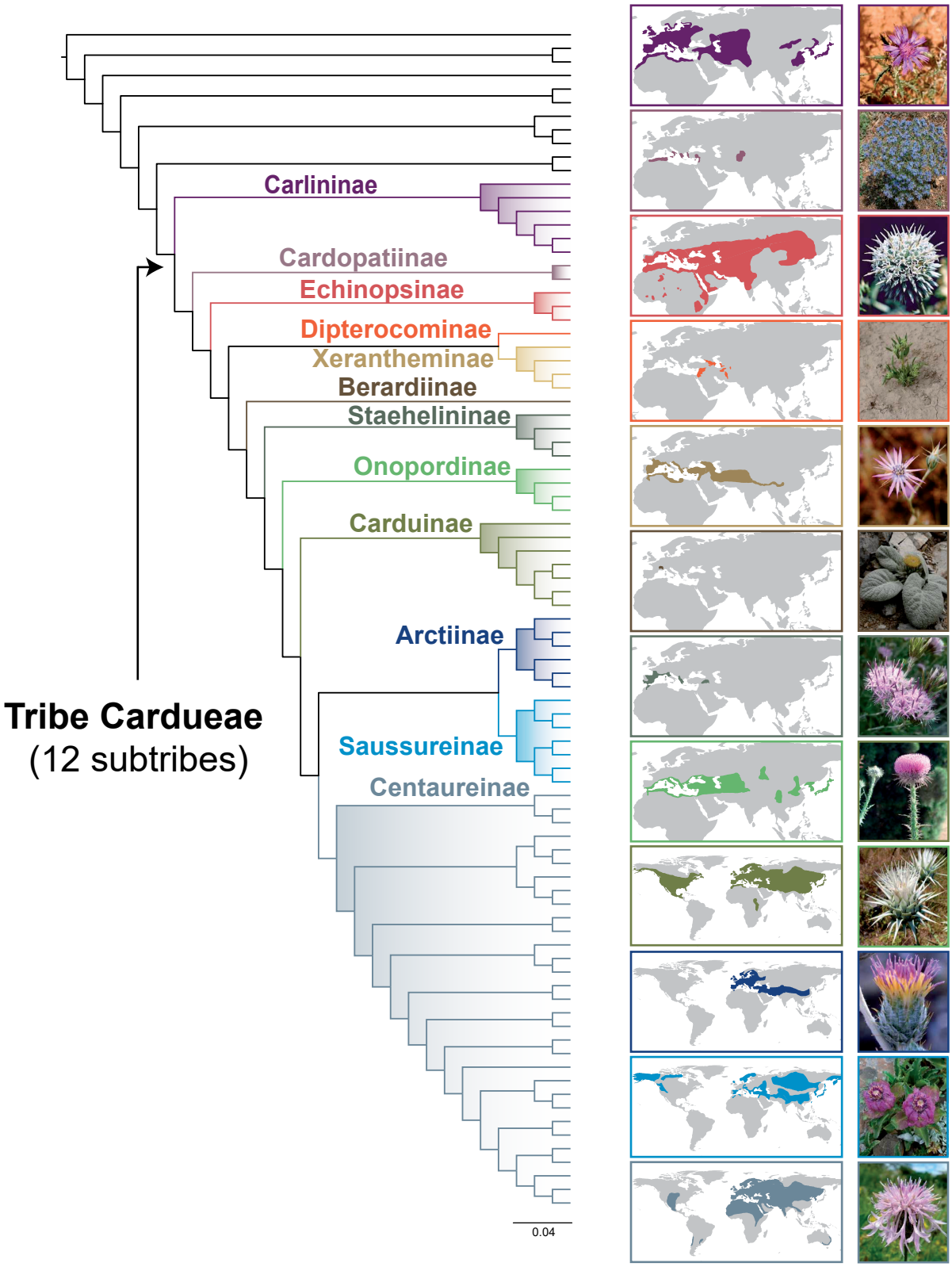
To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Research Data Related to this Submission

There are no linked research data sets for this submission. The following reason is given:
Data will be made available on request

Highlights

- Hyb-Seq resolved with high support relationships among subtribes of Cardueae.
- A new taxonomic treatment of 12 subtribes is proposed.
- The backbone of the Cardueae tree was reconstructed with high resolution.
- Cyto-nuclear phylogenetic discordances are still persisting with NGS data.
- An updated temporal framework for the tribe and subtribes is presented.



Tribe Cardueae
(12 subtribes)

Nuclear and plastid DNA phylogeny of tribe Cardueae (Compositae) with Hyb-Seq data: A new subtribal classification and a temporal diversification framework

Sonia Herrando-Moraira[a,*], Juan Antonio Calleja[b], Mercè Galbany-Casals[b], Núria Garcia-Jacas[a], Jian-Quan Liu[c], Javier López-Alvarado[b], Jordi López-Pujol[a], Jennifer R. Mandel[d], Sergi Massó[a,b], Noemí Montes-Moreno[a], Cristina Roquet[b,e], Llorenç Sáez[b], Alexander Sennikov[f], Alfonso Susanna[a], Roser Vilatersana[a]

[a] Botanic Institute of Barcelona (IBB, CSIC-ICUB), Pg. del Migdia, s.n., ES-08038 Barcelona, Spain

[b] Systematics and Evolution of Vascular Plants (UAB) – Associated Unit to CSIC, Departament de Biologia Animal, Biologia Vegetal i Ecologia, Facultat de Biociències, Universitat Autònoma de Barcelona, ES-08193 Bellaterra, Spain

[c] Key Laboratory for Bio-Resources and Eco-Environment, College of Life Sciences, Sichuan University, Chengdu, China

[d] Department of Biological Sciences, University of Memphis, Memphis, TN 38152, USA

[e] Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA (Laboratoire d'Ecologie Alpine), FR-38000 Grenoble, France

[f] Botanical Museum, Finnish Museum of Natural History, PO Box 7, FI-00014 University of Helsinki, Finland; and Herbarium, Komarov Botanical Institute of Russian Academy of Sciences, Prof. Popov str. 2, 197376 St. Petersburg, Russia

*Corresponding author at: Botanic Institute of Barcelona (IBB, CSIC-ICUB), Pg. del Migdia, s. n., ES-08038 Barcelona, Spain. E-mail address: sonia.herrando@gmail.com (S. Herrando-Moraira).

Abstract

Classification of tribe Cardueae in natural subtribes has always been a challenge due to the lack of support of some critical branches in previous phylogenies based on traditional Sanger markers. With the aim to propose a new subtribal delimitation, we applied a Hyb-Seq approach to a set of 76 Cardueae species representing all subtribes and informal groups defined in the tribe, targeting 1061 nuclear conserved orthology loci (COS) designed for Compositae and obtaining chloroplast coding regions as by-product of off-target reads. For the extraction of the target nuclear data, we used two strategies, PHYLUCE and HybPiper, and 776 and 1055 COS loci were recovered with each of them, respectively. Additionally, 87 chloroplast genes were assembled and annotated. With three datasets, phylogenetic relationships were reconstructed using both concatenation and coalescent approaches. Phylogenetic analyses of the nuclear datasets fully resolved virtually all nodes with very high support. Nuclear and plastid tree topologies are mostly congruent with a very limited number of incongruent nodes. Based on the well-solved phylogenies obtained, we propose a new taxonomic scheme of 12 monophyletic and morphologically consistent

subtribes: Carlininae, Cardopatiinae, Echinopsinae, Dipterocominae (new), Xeranthemeinae (new), Berardiinae (new), Staehelininae (new), Onopordinae (new), Carduinae (redelimited), Arctiinae (new), Saussureinae (new), and Centaureinae. In addition, we further updated the temporal framework for origin and diversification of these subtribes. Our results highlight the power of Hyb-Seq over Sanger sequencing of a few DNA markers in solving phylogenetic relationships of traditionally difficult groups.

Keywords

Asteraceae
COS targets
Subtribes
Systematics
Phylogenomics
Target enrichment

1. Introduction

Cardueae are one of the largest tribes of the 43 described in Compositae (Asteraceae; Funk et al., 2009), with almost 10% of the species of the whole family: 2400 species in 73 genera (Susanna and Garcia-Jacas, 2009). The tribe belongs to subfamily Carduoideae, which is composed of four tribes (Funk et al., 2009; Ortiz et al., 2009): Dicomeae (97 species), Tarchonantheae (13 species), Oldenburgieae (4 species), and Cardueae, the latter representing the 95% of subfamily's diversity (Susanna and Garcia-Jacas, 2009). Carduoideae are a very successful evolutionary lineage within Compositae, which was estimated by Panero and Crozier (2016) to have the second highest diversification rate in the family and a moderate rate of extinction. The number of species is not uniform across the genera of Cardueae. For example, six of them are highly diversified (ca. 200–600 species) and with a high endemism rate (*Carduus*, *Cirsium*, *Centaurea*, *Cousinia*, *Jurinea*, and *Saussurea*), while, on the other extreme, 22 (ca. 30%) are monotypic. Geographically, Cardueae are distributed mainly in the Mediterranean and the Irano-Turanian regions, but they are reported from all continents except the Antarctica. The ecosystems where Cardueae species inhabit are very variable, e.g. Mediterranean landscapes, steppes, semiarid areas, deserts, alpine meadows, or tropical savannahs (Susanna and Garcia-Jacas, 2009). Indeed, in many aspects Cardueae are an extremely heterogeneous group in terms of habit and life-form (perennial, biennial, monocarpic or annual herbs, shrubs, treelets, often spiny), karyology (high variability in chromosome numbers from $x = 6$ to $x = 18$, often with disploidy), or pollen structure (caveate or ecaveate, smooth, scabrate, or spiny). This complexity summed to its high diversity has greatly contributed to the turbulent taxonomic history of Cardueae at several ranks, e.g. tribal and subtribal (see Table 1 for a historical overview) or even misclassifications of some genera in different tribes or subtribes (see examples in Table 2).

A recapitulation of the troublesome history of the systematics of Cardueae was offered by Susanna and Garcia-Jacas (2009). In the first tribal classification of the family, Cassini (1819) divided the present Cardueae into three tribes: Echinopseae, Carlineae, and Cardueae, the latter comprising the subtribes Carduinae and Centaureinae. Years later, Bentham (1873) and Hoffmann (1894) proposed grouping the three former tribes in a single tribe Cardueae, with four subtribes: Carlininae, Echinopsinae ["Echinopinae"], Carduinae, and Centaureinae. This conservative approach was accepted for many years, until Wagenitz (1976) reinstated the tribe Echinopseae. Soon after, Dittrich (1977) returned to Cassini's earlier views, proposing the restoration of tribes Echinopseae, Carlineae, and Cardueae. Bremer (1994) favored the conservative classification of a single tribe Cardueae, which is now generally accepted (Susanna and Garcia-Jacas, 2007, 2009). The only exception is subtribe Cardopatiinae, originally described by Lessing (1832) and only recently recovered 174 years later (Susanna et al., 2006).

Not surprisingly, Carlininae and Echinopsinae have been considered at some time as independent tribes, mainly due to their clear diagnostic characters with respect to the rest of Cardueae assembly (Susanna and Garcia-Jacas, 2009). However, Cardueae are morphologically consistent as a whole entity at tribal level, sharing a unique synapomorphic morphological character within Compositae: style with a papillose-pilose thickening below the branches and the stigmatic areas confined to the inner surface (Susanna and Garcia-Jacas, 2009). Additionally, Cardueae as a tribe has also been broadly confirmed empirically as monophyletic, first by cladistic analyses based on morphology (Bremer, 1987, 1994; Karis et al., 1992), and later by molecular phylogenies (Jansen et al., 1990, 1991; Kim et al., 1992; Susanna et al., 1995, 2006; Garcia-Jacas et al., 2002; Barres et al., 2013).

In the most recent and accepted treatments (Susanna et al., 2006; Susanna and Garcia-Jacas, 2007, 2009), the authors pointed out that four subtribes are natural groups with clear limits (Cardopatiinae, Carlininae, Centaureinae, and Echinopsinae); however, subtribe Carduinae (with total 1700 species, near 70% of the whole tribe diversity) is an unnatural, artificial, and problematic group, which have represented a dumping ground of genera that do not fit morphologically in any of the other subtribes (Susanna and Garcia-Jacas, 2007). The fact that the group is a questionable and heterogeneous assemblage of genera was also reflected in several phylogenetic studies, which have reported the subtribe as paraphyletic because Centaureinae are nested on Carduinae (Susanna et al., 1995, 2006; Häffner and Hellwig, 1999; Garcia-Jacas et al., 2002; Barres et al., 2013). The alternate solution of combining subtribes Carduinae and Centaureinae in one enormous subtribe was discarded, owing to the impractical constitution of a huge subtribe encompassing 2300 species and 90% of the species of the whole tribe (Susanna and Garcia-Jacas, 2007).

With the intention of addressing the unresolved diversity within subtribe Carduinae, some informal morphological groups were described within it (Susanna and Garcia-Jacas, 2007), which could be considered subtribes for a more natural classification: "[the alternative is] splitting present Carduinae in at least seven new subtribes (many of them presently unsupported): Xeranthemineae, Staehelininae, Berardininae, Onopordinae, Carduinae, Arctiinae, and Saussureinae […]" (Susanna and Garcia-Jacas, 2009). In view of the lack of support for the monophyly of all segregate subtribes, especially for the most important in terms of species number (core Carduinae, Arctiinae, and Saussureinae), no formal proposal has previously been performed. A well-resolved and supported phylogenetic hypothesis for the major

lineages within Carduinae is still lacking, and it would be very useful to confirm the morphological alliances or generic complexes proposed by Susanna and Garcia-Jacas (2009).

An additional taxonomic problem within Cardueae is that some genera or complex of genera have been classified within one subtribe or another depending on the classification followed (see some of the cases in Table 2). Moreover, in extreme cases, some genera have been independently classified in Cardueae or in other tribes of Compositae (e.g. *Berardia* in tribe Mutisiae, or *Dipterocome* in tribe Calenduleae). A great proportion of subtribal misplacements were reported for uncertain cases between subtribes Carduinae and Centaureinae, which could be attributed to the inconspicuous morphological differences between both subtribes (based on microcharacters of achene and pappus) that are sometimes lacking in herbarium specimens, or are immature structures during field collections (Susanna et al., 2006). Even though many cases have been successfully resolved with the aid of the molecular phylogenies obtained with Sanger sequencing (see references in Table 2), more changes in the adscription of genera could occur in a new subtribal framework for Cardueae.

To date, all previous molecular studies of Cardueae have been based on Sanger sequencing data (Susanna et al., 1995, 2006; Häffner and Hellwig, 1999; Garcia-Jacas et al., 2002; Barres et al., 2013), with the largest dataset constructed with four chloroplast regions (*trnL-trnF*, *matK*, *ndhF*, *rbcL*) and one nuclear ribosomal marker (ITS; Barres et al., 2013). Although subtribal clades of Carlininae, Echinopsinae, Cardopatiinae, and Centaureinae are well supported, Carduinae remains paraphyletic when Centaureinae are removed, and not all the informal groups have been recovered as monophyletic. Moreover, relationships among subtribes have not been resolved, and the backbone of the phylogenetic tree remains undefined without a supported dichotomous bifurcating pattern. Accordingly, the available divergence time estimation of Cardueae is based on a partially resolved phylogenetic tree resulting from the analysis of a combined dataset of chloroplast markers (Barres et al., 2013). This dating could be significantly improved in terms of the methodological approach and the sequence data used.

In the last decade, next generation sequencing (NGS) has emerged as an important methodological advance for solving phylogenetic problems (see the review of Harrison and Kidner, 2011). Plant phylogenies of historical taxonomically complex groups are becoming resolved at different taxonomic levels, e.g. order (Ranunculales, cf. Lane et al., 2018; or Zingiberales, cf. Carlsen et al., 2018), family (Goodeniaceae, cf. Gardner et al., 2016), tribe (Shoreeae, cf. Heckenhauer et al., 2018), or species complexes as *Claytonia* (Stoughton et al., 2018) or *Amaranthus* (Viljoen et al., 2018). For Compositae, Mandel et al. (2014) designed a probe set that hybridizes with 1061 nuclear conserved orthology loci (hereafter COS), which in combination with genome skimming allowed also to recover other parts of the genome such as chloroplast regions (defined as Hyb-Seq technique; Weitemier et al., 2014). This methodological workflow has proved successful in Compositae-wide studies (Mandel et al., 2015, 2017) and was tested in a recent research focused on highly radiated genera within Cardueae (Herrando-Moraira et al., 2018). However, the COS locus set has not been yet applied to taxonomic delimitation within the family.

In this study, we apply the Hyb-Seq method to a sample of 76 species representing all subtribes and informal suprageneric groups of Cardueae with the main goals to: (1) obtain a well-defined phylogeny of the high-level lineages in the tribe with nuclear and chloroplast data; (2) propose a new subtribal

classification, especially focused on testing the splitting of subtribe Carduinae into smaller and more practical natural subtribes; (3) examine previously unresolved phylogenetic relationships between subtribal lineages at the backbone of the tree; and (4) update the temporal framework of the tribe and subtribes origin and diversification.

2. Materials and methods

2.1. Taxon sampling

To obtain a complete sampling of Cardueae, we included representatives of all described subtribes based on the taxonomic treatment of Susanna and Garcia-Jacas (2009): (1) Carlininae, 6 species; (2) Cardopatinae, 2 species; (3) Echinopsinae, 3 species; (4) Carduinae, 34 species; and (5) Centaureinae, 31 species. The sampling strategy was based on maximizing the number of species of unresolved informal groups, especially within Carduinae and Centaureinae, and was proportional to the total number of genera and species classified in each group. In total, 76 different species of Cardueae were included. We did not include all genera within the tribe because the adscription of many of those with taxonomic problems had already been confirmed in previous works (Table 2). In addition, we incorporated 5 species as representatives of subfamily Carduoideae outside of Cardueae; 5 species from other subfamilies within Compositae (2 from Mutisioideae, 2 from Barnadesioideae, 1 from Famatinanthoideae); and finally, *Nastanthus patagonicus* from Calyceraceae, which has been recognized as the sister family to Compositae (Mandel et al., 2017). Overall, 58 species were newly sequenced for this study, and the data of the remaining 29 were directly obtained from raw reads from Mandel et al. (2014, 2017) or Herrando-Moraira et al. (2018). See Supplementary Table S1 for details of each sampled species.

2.2. DNA extraction, library and capture preparation, and sequencing

Approximately 10–30 mg of dried plant material per sample was weighed and homogenized with a Mixer Mill MM 301 (Retsch®, Haan, Germany). The DNA was extracted with the DNeasy plant mini kit (Qiagen, Valencia, CA, USA) or the E.N.Z.A SP Plant DNA Mini Kit (Omega Bio-Tek Inc., Norcross, GA, USA) according to the manufacturer's protocol. The total genomic DNA quantity was measured with the Qubit™ 3.0 Fluorometer (Thermo Scientific, Waltham, MA, USA). The standardized DNA (1 µg in 70 µl) was sheared in the Genomics Unit of the Centre for Genomic Regulation (CRG, Barcelona, Spain) using a Covaris S2 System (Covaris, Woburn, MA, USA) in microTUBEs with a sample volume of 50 µl and a target peak set to 400 bp. Sequencing libraries and subsequent sequence capture were conducted as specified in Herrando-Moraira et al. (2018). Additionally, for the species newly sequenced for this study (see Supplementary Table S1), we conducted a library spiking with the following proportions: 40% of unenriched solution and 60% of enriched libraries. The final spiked library pools were sequenced (pair-end 100 bp) in the DNA Sequencing Core CGRC/ICBR of the University of Florida on one lane of an Illumina HiSeq 3000 (Illumina, USA).

2.3. Raw data processing

Demultiplexed raw sequence reads were checked with FastQC v.0.10.1 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) in order to perform a first quality control step. Subsequently, raw reads were trimmed by a quality based assessment (sliding-window set to 5:20), and were also subjected to an adapter trimming using Trimmomatic v.0.36 (Bolger et al., 2014). In addition, the software was programmed to remove those cleaned reads that after trimming were less than 36 bp, and those with a missing forward or reverse pair.

2.4. Nuclear data extraction

Following the methodological workflow proposed in Herrando-Moraira et al. (2018), the 1061 target COS loci were extracted with two orthology-detection pipeline packages: PHYLUCE v.1.5 (Faircloth, 2015) and HybPiper v.1.1 (Johnson et al., 2016). We used both approaches to be sure that the distinct paralog processing strategy was not affecting phylogenetic reconstructions in terms of topology and branch support values (Herrando-Moraira et al., 2018).

For the PHYLUCE method, the trimmed reads were *de novo* assembled into contigs with the software SPAdes v.3.9.0 (Bankevich et al., 2012), specifying different predefined k-mer lengths (21, 33, 55, and 77). Then, the recovered contigs were mapped back to target references using LASTZ (Harris, 2007), and were extracted following the methodological specifications detailed in Herrando-Moraira et al. (2018). For the HybPiper method, the trimmed reads were firstly mapped to the targets using BWA (Li and Durbin, 2009), and secondly were assembled into contigs also using SPAdes.

The multi-fasta files obtained for each target locus both with PHYLUCE and HybPiper were aligned with MAFFT v.7.266 (Katoh and Standley, 2013) using the *auto* setting mode. The ambiguously aligned regions were removed with trimAl v.14 (Capella-Gutiérrez et al., 2009) applying the *automated1* flag. The loci extracted for less than three species were removed, in the case of PHYLUCE 126 from the total of 902 loci, and in HybPiper 1 locus from the total 1057 loci. One additional locus was removed from the HybPiper dataset due to its short length recovered after the alignment trimming (3 bp). Finally, for each target extraction method, two datasets were constructed in order to perform two different phylogenetic inference analyses. One consisted of the trimmed aligned sequences of each locus separately, and the other consisted of a supermatrix obtained by concatenating all trimmed aligned loci with FASconCAT-G v.1.02 (Kück and Longo, 2014). The summary statistics of the supermatrices were calculated with AMAS (Borowiec, 2016).

2.5. Chloroplast data extraction

For the reconstruction of the chloroplast genomes, we used the off-target reads also recovered with the Hyb-Seq approach. The chloroplast extraction was conducted using the pipeline package MITObim v.1.9 (Hahn et al., 2013) with the default conditions. The package has incorporated the module MIRA v.4.0.2 (Chevreux et al., 1999), which is used in mapping mode. In the first step, it identifies the more conserved

regions between the total readpool and a phylogenetically related reference genome (in this case, *Cirsium arvense* NCBI accession number NC_036965.1) by mapping the trimmed interleaved reads to this initial reference. Then, the mapped reads are assembled into contigs and a new reference sequence or bait is created. To extend the reference and close the gaps, two more steps are performed. The second consists of separating from the total read pool the reads that overlap with the reference, and the third recovers the reads separated to map them back to the bait, thus building a new more extended reference sequence through an assembly process of new mapped reads. These two last steps are iteratively repeated until a stationary stage of the mapped and assembled reads is reached (Hahn et al., 2013).

The chloroplast genomes recovered were annotated with the web tool application GeSeq (Tillich et al., 2017), which uses a customizable reference database to annotate the genomes using BLAT-driven best-match approach. As a database, we selected all the available annotations in NCBI of Cardueae representatives (*Carthamus tinctorius* NC_030783.1, *Centaurea diffusa* NC_024286.1, *Cirsium arvense* NC_036965.1, *Cirsium eriophorum* NC_036966.1, *Cirsium vulgare* NC_036967.1, *Cynara baetica* NC_028005.1, *Cynara cornigera* NC_028006.1, *Cynara humilis* NC_027113.1, *Saussurea chabyoungsanica* NC_036677.1, *Saussurea involucrata* NC_029465.1, *Saussurea polylephis* NC_036490.1, and *Silybum marianum* NC_028027.1). To extract the coding regions of genes (CDS) in separate files from the global multi-fasta matrix generated with GeSeq, we used the script "Phyluce_assembly_explode_get_fastas_file" of PHYLUCE package with slight modifications. The recovered CDS were individually aligned with the MACSE codon aligner (Ranwez et al., 2011). The alignments were visualized in SeaView v.4.7 (Gouy et al., 2009), and 10 species from the initial sampling were excluded due to the high presence of frameshifts and preceding codon stops. To remove the poorly aligned regions from the alignments, we used the program Gblocks v.0.91b (Castresana, 2000) with the option "codon" to trim entire codon sets. A total of 87 CDS regions were recovered. We concatenated all CDS in a single supermatrix (chloroplast dataset) with FASconCAT-G v.1.02, assuming that the chloroplast genome is considered to be haploid, nonrecombinant, generally uniparentally inherited, and "single copy" (Small et al., 1998, 2004).

2.6. Phylogenetic inference analyses

The phylogenetic reconstruction of Cardueae was performed under two different approaches for nuclear data: the concatenation approach (using the supermatrix dataset) and the coalescence approach (using the individual gene trees for inferring the species tree). For chloroplast data, the phylogenetic tree inference was only performed under the concatenation approach.

For the concatenation approach, Maximum Likelihood (ML) analyses were run using the software RAxML v.8.2.9 (Stamatakis, 2014), which were conducted on XSEDE in the CIPRES Science Gateway v.3.1 (Miller et al., 2010). We selected the algorithm of simultaneously searching the best ML tree (from 10 randomized maximum parsimony starting trees) and performing a rapid bootstrapping (1000 replicates). Branches were considered as statistically supported for bootstrap (BS) values > 70% (Hillis and Bull, 1993). In relation to the partition scheme, each locus (in the nuclear datasets) or gene (in the chloroplast dataset)

was considered a unit of partition, using the evolution model of GTRGAMMA following recommendations of Stamatakis (2006). Output trees were visualized and exported in FigTree v.1.4.3 (Rambaut, 2016).

For the coalescence approach, gene trees were inferred also using RAxML, under the same conditions as for the concatenation based analyses but with a bootstrap resampling of 200 replicates. Resulting unrooted gene trees were inputted into ASTRAL-III v.5.5.3 (Zhang et al., 2018) to infer the species tree. The support values were calculated using local posterior probabilities (LPP; Sayyari and Mirarab, 2016). Branches with LPP > 0.95 were considered as strongly supported.

2.7. Gene trees concordance analyses

Topological conflicts among gene trees and the species tree were explored using the software Phyparts (Smith et al., 2015) on the nuclear HybPyper dataset. This methodological approach consists of mapping, for each node or bipartition of interest, the level of concordance/discordance between the different individual gene trees on supporting the reference tree topology, in our case the ASTRAL species tree. First, gene and species trees were rooted with the online tool STRAW (Shaw et al., 2013). The outgroup selection was based on the species divergence rank recovered by the species ML tree inferred under concatenation. When the first preferred outgroup was missing, STRAW searches the next specified species. Then, Phyparts was run, considering only the branches of gene trees with more than 33% BS support. Finally, the script phypartspiecharts.py (https://github.com/mossmatters/MJPythonNotebooks) was used to summarize and map the Phyparts output on the ASTRAL species tree. Owing to computing limitations of Phyparts, we could not run the analysis for the entire set of 1055 nuclear gene trees of HybPiper dataset, therefore a subset of 750 random gene trees was selected.

2.8. Divergence time analysis

The divergence time analysis was performed on the phylogenetic tree obtained under concatenation from the nuclear HybPiper dataset. This tree was time-calibrated rescaling the branch lengths using the penalized likelihood method (Sanderson, 2002), implemented in the software treePL (Smith and O'Meara, 2012). This approach yields similar results than the software BEAST (e.g. Lagomarsino et al., 2016; Stubbs et al., 2018), but it runs faster on larger datasets. The dating procedure was divided into two main stages, which consisted on: (1) selection of the optimal model parameters; and (2) running the analysis with the optimal parameters selected and, additionally, accounting for the uncertainty in calibration points to obtain confidence intervals (95% CI) in the estimated node ages.

For both stages, five calibration points (CP) were used as node age constrains. One was a secondary dated node corresponding to the origin of Compositae family (69.56 Myr with a 95% CI 59.02–80.17; coded here as CP1) reported in Panero and Crozier (2016). The other four points were based on fossil records: (CP2) the capitulescence of *Raiguenrayun cura* that was dated at 47.5 Myr (Barreda et al., 2012), which was used to constrain the clade of subfamily Mutisioideae + subfamily Carduoideae as in the most recent interpretation by Panero and Crozier (2016), instead of placing it at the crown Compositae as in Barres et al. (2013); (CP3) the achenes identified as belonging to *Cirsium* with an age of 14 Myr (Mai,

1995), which were placed at the stem node of the *Carduus-Cirsium* clade following Barres et al. (2013); (CP4) the achenes assigned to *Arctium* (López-Vinyallonga et al., 2009) and dated at 8 Myr (Mai, 2001), which were placed at the stem node of *Arctium*, correcting the misplacement of this fossil in Barres et al. (2013), where it was used at the split of *A. minus* and *A. lappa*; and (CP5) the pollen of *Centaurea* type *Cyanus* dated at 6 Myr (Wagenitz, 1955; Ivanov et al., 2007), which was placed at the stem node of *Centaurea* subgenus *Cyanus*. This is a new, more precise and better resolved *Centaurea* fossil that the pollen and achene records of Mai (1995) and Popescu (2002) that were used in Barres et al. (2013), where they were placed at stem age of *Centaurea* due to their uncertain placement within the genus.

In the first dating stage, an initial "priming" run (prime command) was carried out to detect the optimal parameter settings (opt = 4, optad = 2, and optcvad = 2). Subsequently, the best smoothing rate was also evaluated through a random subsample and replicate cross-validation procedure (selecting thorough and randomcv commands), allowing varying it from 0.001 (cvstart) to 1000 (cvstop). The best smoothing value resulted in 0.1 after cross-validation analysis, which was selected from the lowest value of a Chi-Square test. The two runs of the first stage were run using the range of 95% CI value for the CP1 (minimum = 59.02 and maximum = 80.17) and the lower bound of fossil age estimations (minimum = fossil age) for CP2-CP5.

In the second stage, we used our own R (R Core Team, 2014) script (Appendix A) to obtain 10,000 random values for each CP in order to account for the uncertainty of the node age estimations. Specifically, the random values were generated under a normal distribution for CP1 (based on 95% CI estimations of Panero and Crozier, 2016; mean = 69.56 and SD = 3) and a lognormal distribution for CP2-CP5 (mean = fossil age and SD = 1.1). Once values were extracted, a new data frame was created in a spreadsheet (Appendix B), which was then used to randomly select 100 unique combinations of five numbers, i.e. one for each CP from the random pool of 10,000. Finally, 100 independent treePL analyses, each one with CP constrained to a single value (minimum = maximum), were ran using the same phylogenetic input tree and the optimized parameters found in the first stage. The resultant 100 dated tree files were modified to fit the format of TreeAnnotator v.1.7.5 (Drummond et al., 2012), and were introduced in this software to obtain a maximum clade credibility tree chronogram with median node heights and corresponding CIs.

3. Results and discussion

3.1. Performance of Hyb-Seq at subtribal level within tribe Cardueae

The Hyb-Seq NGS technique designed for Compositae (Mandel et al., 2014) is confirmed here as a powerful tool to reconstruct highly resolved phylogenies at deep taxonomic levels, in line with previous studies more focused on the methods (Mandel et al., 2015, 2017; Herrando-Moraira et al., 2018). As far as we know, this is the first study in Compositae that uses the COS loci workflow to reach taxonomic conclusions, specifically aimed to clarify the entangled delimitation of subtribes within Cardueae. The transition from traditional Sanger datasets, up to 5 markers, to NGS datasets, with 1142 markers (Table 3),

has resulted in obtaining the most helpful dichotomous and confidently supported divergence pattern reported to date for Cardueae (Figs. 1, 2, and 3, Supplementary Figs. S1 and S2).

On average 5,420,504 reads (SD = ±3,461,361) were sequenced per species (Supplementary Table S2). From the total 1061 COS loci, 776 were finally recovered with PHYLUCE and 1055 with HybPiper, which resulted in matrices of trimmed aligned sequences of 492,549 bp (189,716 parsimony informative [PI] sites) and 332,260 bp (123,731 PI sites), respectively (Table 4). When parameters related to missing data were compared between PHYLUCE and HybPiper (see parameters 4 and 11 in Table 4), it was remarkable that PHYLUCE is less efficient than HybPiper in the sequence extraction process from on-target reads, as documented in Herrando-Moraira et al. (2018). However, phylogenetic trees did not differ in topology or branch supports between both extraction methods at subtribal level (Figs. 1 and 3, Supplementary Figs. S1 and S2), even though PHYLUCE uses a more restrictive procedure to remove potential paralogs than HybPiper (see Fig. 1 in Herrando-Moraira et al., 2018). As no differences were observed, results from HybPiper method were analyzed in deep to support taxonomic proposals. Interestingly, this lack of differences could be reflecting that paralogy incidence is relatively low in Cardueae. In agreement, the study of Jones et al. (in prep.), which tests the COS loci applicability for seven tribes of Compositae, found that Cardueae was the tribe with the lowest number of paralogs (average = 140, similar to the value found here, average = 130), in compared to others like Vernonieae (average = 256). Also in concordance, Huang et al. (2016) did not detect any recent whole genome duplication event in Cardueae, in contrast to other Compositae tribes, which could involve paralogy issues.

As other studies have shown, Hyb-Seq can provide nearly complete datasets of chloroplast genomes (e.g. Weitemier et al., 2014; Folk et al., 2015; Carlsen et al., 2018). Regarding the present study, we were able to retrieve 87 protein coding regions of the chloroplast genome from off-target reads, which after concatenation resulted in a matrix of 78,531 bp (4290 PI sites; Table 4). Indeed, the increase of phylogenetically informative characters is responsible of notably higher topological resolution and clade support in comparison with the previous results reported with Sanger sequencing data for Cardueae (Susanna et al., 1995, 2006; Häffner and Hellwig, 1999; Garcia-Jacas et al., 2002; Barres et al., 2013; Table 3). In a similar way, this effect has been already documented for other plant groups (e.g. *Pinus*, Parks et al., 2009; *Inga*, Nicholls et al., 2015), which have benefited tremendously from the advent of high-throughput sequencing techniques.

3.2. The new subtribal classification of tribe Cardueae

As it could be expected with our sampling, we recovered a monophyletic pattern for subfamily Carduoideae as currently defined (tribes Dicomeae, Oldenburgieae, Tarchonantheae, and Cardueae; cf. Funk et al., 2009). Tribe Cardueae results again a robust taxonomic unit, in agreement with former studies based on morphology and molecular data (Bremer, 1987, 1994; Jansen et al., 1990, 1991; Karis et al., 1992; Kim et al., 1992; Susanna et al., 1995, 2006; Garcia-Jacas et al., 2002; Barres et al., 2013). Specifically, the clade of Cardueae presented high support values in all the analyses (Figs. 1, 2, and 3). The alternative option of rising up some subtribes to a tribal rank (see Table 1) is once again discarded. As suggested by Susanna et al. (2006), it is unpractical to fragment a natural group that is strongly phylogenetically consistent and

can be easily recognized by macromorphological characters. In this way, the classification criteria for the definition of Compositae tribes are also maintained homogeneously along the family (Funk et al., 2009).

At subtribal level, the high resolution of phylogenies provided by the Hyb-Seq data makes possible the proposal of a new classification. Subtribal limits established here are based on the application of the integrative taxonomy principles (Dayrat, 2005; Schlick-Steiner et al., 2010): firstly, considering the morphological entities to be recognized (Susanna and Garcia-Jacas, 2007, 2009), and then verifying their monophyly with molecular data. Following this model, a total of 12 subtribes are suggested. Five of them had been already previously recognized and are also confirmed with our results (Carlininae, Echinopsinae, Cardopatiinae, Carduinae, and Centaureinae; Susanna et al., 2006); seven new subtribes result from partitioning the former wide and paraphyletic subtribe Carduinae into new monophyletic subtribes based on informal morphological groups (Dipterocominae, Xeraantheminae, Berardiinae, Staehelininae, Onopordinae, Arctiinae, and Saussureinae) most of them outlined by Susanna and Garcia-Jacas (2009). These new subtribes are highly heterogeneous in terms of species richness, encompassing from relict single lineages to highly recently radiated groups. For instance, Dipterocominae and Berardiinae are defined as monotypic subtribes, comprising only *Dipterocome pusilla* and *Berardia subacaulis*, respectively. On the other hand, Arctiinae and Saussureinae harbor ca. 600 and 550 species, respectively (Susanna and Garcia-Jacas, 2007). For a complete description of subtribes see below *Taxonomic proposal*. The new phylogenetic reconstruction of high-level lineages presented here will be used as a classification basis for future studies on the systematics of Cardueae.

Certainly, the historical overview of the tribe (Table 1) reveals that Carlininae, Echinopsinae, and Centaureinae are the best morpho-molecular defined entities, which have been long recovered with statistical support since the first phylogenetic studies (Table 3). In accordance, we obtained high clade supports for these subtribes (BS = 100 and LPP = 1; Figs. 1, 2, and 3). Carlininae initially showed a monophyletic pattern in Garcia-Jacas et al. (2002). However, the incorporation of *Tugarinovia mongolica* broke the monophyly of the subtribe in Susanna et al. (2006). This was probably caused by the fact that it is a dioecious species, a trait unique in the tribe (Susanna and Garcia-Jacas, 2009) and relatively uncommon in angiosperms (6% of species; Renner and Ricklefs, 1995). It has been suggested that dioecy has the negative impact of decreasing diversification rates (Heilbuth, 2000; Kay et al., 2006) with consequent phylogenetic misplacements (Vamosi et al., 2003). In Barres et al. (2013) and in our present study, the increase in the number of molecular markers aided to recover the monophyly of Carlininae, including *Tugarinovia mongolica* within this subtribe as a sister lineage to the rest of members. Echinopsinae presented a persistent monophyly across all molecular studies (Table 3), except for the chloroplast-based phylogeny of Garcia-Jacas et al. (2002), in which the limited sequence variation found in chloroplast *mat*K gene could have been the responsible for the low branch support of Echinopsinae. Similarly, the monophyly of subtribe Centaureinae has been fully supported in former phylogenetic studies (Table 3), with the exception of the chloroplast tree obtained with the *mat*K gene by Garcia-Jacas et al. (2002).

Although Cardopatiinae are a relatively newly recovered subtribe (Susanna et al., 2006), it has resulted also in a strong-supported clade in recent phylogenies with both nuclear and chloroplast sequence data (Table 3), besides its morphological singularity (Susanna and Garcia-Jacas, 2007, 2009). In agreement, we can confirm here that the two species of Cardopatiinae (*Cardopatium corymbosum* and *Cousiniopsis*

437 *atractyloides*) constitute a well-defined evolutionary segregation (BS = 100 and LPP = 1; Figs. 1, 2, and 3),
438 therefore its taxonomic rank as subtribe seems appropriate.
439 Conversely, the new seven subtribes have not always been well delineated in Sanger-based
440 phylogenies, at least with enough confidence to propose a formal subtribal splitting (see Table 3). From the
441 morphological informal groups of Carduinae described in Susanna and Garcia-Jacas (2009), the
442 *Xeranthemum* group is one of the best supported of all datasets and analyses (BS = 100 and LPP = 1; Figs.
443 1, 2, and 3). Thus, its consideration as subtribe Xerantheminae can be corroborated. Our results are not
444 surprising, because since the first phylogenies the Xerantheminae species were recovered as a monophyletic
445 and statistically highly supported group, both with nuclear and chloroplast data (Table 3).
446 One unexpected finding was found in the former *Xeranthemum* group, which is the phylogenetic
447 isolated position of *Dipterocome pusilla*. This species had not been classified in Cardueae until the study
448 of Anderberg et al. (2007), confirmed by Barres et al. (2013), in which it was clustered with the rest of
449 Xerantheminae members in the combined cpDNA (*mat*K, *ndh*F, *rbc*L, and *trn*L-*trn*F) tree, but without
450 support in a phylogeny of the ITS nuclear region. However, in the present study we found that *Dipterocome*
451 is placed in a different clade from Xerantheminae in four of the five reconstructed phylogenies (Figs. 2 and
452 3, Supplementary Figs. S1 and S2). Only in the tree reconstructed under concatenation with the nuclear
453 HybPiper dataset *Dipterocome* was grouped with Xerantheminae (BS = 97; Fig. 1), in agreement with the
454 results of Barres et al. (2013). This may be caused by the bias produced by concatenation analysis of NGS
455 datasets with a large number of loci (Kubatko and Degnan, 2007; Edwards et al., 2016), which can result
456 in long branch attraction artifacts (Liu et al., 2015). Morphologically, Susanna and Garcia-Jacas (2009)
457 pointed out that floral characters of *Dipterocome* show close relationship with the *Xeranthemum* complex.
458 Nevertheless, a histological analysis of the capitula structure of *Dipterocome* has shown that the purported
459 "florets" of *Dipterocome* are actually one-flowered capitula (Susanna, unpubl. data) grouped in second-
460 order synflorescences, an extremely infrequent structure in the tribe shared only by the genus *Echinops*.
461 Overall, it seems that there is enough molecular and morphological evidence to classify *Dipterocome* in a
462 separate monotypic subtribe.
463 Two other controversial genera considered to belong to the former Carduinae are *Berardia* and
464 *Staehelina*. Although both genera do not present obvious morphological affinities (Susanna and Garcia-
465 Jacas, 2009), some molecular phylogenies suggested that they could be evolutionary closely related. The
466 genus *Staehelina* has long been reported as a genetically isolated group within the tribe, always showing a
467 monophyletic pattern throughout all phylogenetic reconstructions (Table 3). Regarding its relationships
468 with *Berardia*, the combined nuclear plus chloroplast tree in Susanna et al. (2006) grouped both genera in
469 an unsupported clade; but only considering nuclear data (ITS), they appeared not directly linked but as
470 adjacent lineages as detected in nuclear and chloroplast phylogenies of Barres et al. (2013). Here, we also
471 identified the same two phylogenetic signals. Even though, only one of five phylogenetic analyses
472 (chloroplast dataset analyzed under the concatenation approach) resulted in a moderately-high grouping of
473 both genera (BS = 76; Fig. 2), and the rest separated *Berardia* and *Staehelina* (Figs. 1 and 3, Supplementary
474 Figs. S1 and S2). As we hypothesized for the case of *Dipterocome*, *Berardia* could be misplaced in some
475 chloroplast phylogenies by a long branch attraction effect in concatenated analyses due to its relict character

476 without any known alive direct congener. Conclusively, *Berardia* and *Staehelina* could be considered as
477 separate new subtribes Berardiinae and Staehelininae based on the new molecular evidence presented.
478 The *Onopordum* group is another lineage of former Carduinae with a consistent morpho-genetic
479 historical division (Susanna and Garcia-Jacas, 2009). Our genomic phylogenies provide evidence for the
480 monophyly of the *Onopordum* group (BS = 100 and LPP = 1; Figs. 1, 2, and 3) as other previous studies
481 (Table 3), except in Garcia-Jacas et al. (2002) in which the phylogenetic resolution power of the chloroplast
482 *mat*K region seems limited as mentioned above for other groups. Hence, it seems reasonably to assign the
483 subtribal rank as Onopordinae.
484 One of the largest and most difficult complexes to delimit in former Carduinae is the *Carduus-Cirsium*
485 group (ca. 500 species; Susanna et al., 2006). Its complicate delimitation was mainly attributed by Susanna
486 et al. (2006) to the fact that the group encompasses a remarkable variety of life strategies; annual or biennial
487 species (e.g. *Carduus*, *Galactites*, *Picnomon*, *Silybum*, or *Tyrimnus*) or perennials (as many *Cirsium*,
488 *Cynara*, *Lamyropsis*, or *Ptilostemon*). Certainly, different generation times that affect mutation rates may
489 hinder a proper phylogenetic comparison among species as detected in some cases in Cardueae (see López-
490 Vinyallonga et al., 2009 and references therein). The overview of previous phylogenetic studies reflects the
491 difficulty of reconstructing the *Carduus* group as monophyletic, for which additional sampling is necessary
492 to resolve generic delimitations among these genera. The monophyly of the *Carduus* group was only
493 achieved in three from the total eight Sanger-based phylogenies (Table 3). In contrast, we found a strong
494 phylogenetic signal with Hyb-Seq data (BS = 92–100 and LPP = 1; Figs. 1, 2, and 3). This supports the
495 morphological hypothesis of Susanna et al. (2006) and subtribal delimitation of Carduinae. Susanna and
496 Garcia-Jacas (2009) suggested a posterior division of the *Carduus* group into *Cynara* and *Carduus-Cirsium*
497 complexes, but our phylogenies did not show a strictly bifurcating divergence pattern for these complexes
498 (Figs. 1, 2, and 3). Thus, we would reject to formalize their segregation at subtribal rank.
499 The *Arctium-Cousinia* and *Jurinea-Saussurea* complexes are other classically entangled groups of
500 former Carduinae. Although both conform well defined assemblies by morphology (Susanna and Garcia-
501 Jacas, 2009), there has not been any Sanger-based phylogeny able to correctly recover the morphological
502 pattern as monophyletic lineages (Table 3). Only two isolated molecular studies, Häffner and Hellwig
503 (1999) and Susanna et al. (2006), reported *Arctium-Cousinia* and *Jurinea-Saussurea* as monophyletic,
504 respectively. In recent times, the application of Hyb-Seq technique and the COS loci set of Mandel et al.
505 (2014) for Compositae, aided to shed light on phylogenetic connections of these four genera. Phylogenomic
506 trees performed in Herrando-Moraira et al. (2018) showed a clear sister relationship between *Arctium-*
507 *Cousinia* and *Jurinea-Saussurea*, which also formed two highly supported clades. In agreement, on a wider
508 taxonomic sampling of the tribe, we found the same monophyletic pattern for both complexes under
509 concatenation and coalescence approaches with nuclear data (BS = 100 and LPP = 1; Fig. 1 and 3).
510 However, the chloroplast phylogenies analyzed under concatenation recovered an unsupported clade of
511 *Saussurea-Cousinia-Arctium* (Fig. 2) with *Jurinea* as a separate lineage, hence leaving the complex
512 *Jurinea-Saussurea* as a paraphyletic assembly (Fig. 2). Similarly, a morphologically incongruent grouping
513 pattern has been also recovered in the Sanger-based chloroplast phylogenies of Barres et al (2013), where
514 the genera *Saussurea-Jurinea-Cousinia* were clustered together with support and *Arctium* was placed as a
515 sister clade.

Overall, it seems that even with massive Hyb-Seq data, a clear cyto-nuclear discordance exists between evolutionary histories inferred for *Arctium-Cousinia* and *Jurinea-Saussurea* groups. Two factors could explain such pattern of incongruent genealogies (Morales-Briones et al., 2018), one more related with tree inference methods (phylogenetic error), and the other with the inherent evolutionary history of the genera (incomplete lineage sorting [ILS] or hybridization). Firstly, it is stated that using gene fragments with limited phylogenetic signal may derive in an increase of both stochastic and systematic errors in reconstructed phylogenies (Lemmon et al., 2009, Lemmon and Lemmon, 2013; Carlsen et al., 2018). Hence, if we compare our nuclear vs. chloroplast datasets, a large difference exists in terms of phylogenetic informativeness, for instance, in PI sites (ca. 38% vs. 5.5%; Table 4) or the average branch length of trees (0.04 vs. 0.004). Chloroplast data from protein coding regions, in comparison with hundreds of nuclear loci, could be providing limited phylogenetic information, which may explain the differences in the resulting tree topologies.

Secondly, ILS or reticulation events may explain conflicting topologies and gene tree incongruences (García et al., 2017). The short internal branches observed at the base of genera in chloroplast tree from concatenation analysis (Fig. 2B) could match the signals left by ILS, i.e. a scenario of a rapid diversification of genera in a short timeframe (rapid bifurcating cladogenesis) or a simultaneous polychotomous divergence (hard polytomy). Past hybridization events could also have taken place as suggested by the concatenation chloroplast tree topology (Fig. 2). The chloroplast genome of the ancestral *Saussurea* lineage could have been lost due to an organelle capture (Stegemann et al., 2012), thus, the currently observed topology is that retained from an ancestral hybridization with the *Arctium-Cousinia* lineage. Cyto-nuclear discordances due to introgressive hybridization have been suggested in other plant groups like *Ficus* through chloroplast NGS phylogenies (Bruun-Lund et al., 2017). However, more explicit methods to address these processes should be performed (e.g. Folk et al., 2017; García et al., 2017; Mitchell et al., 2017; Knowles et al., 2018; Morales-Briones et al., 2018), in addition to using a wider species sampling.

Taking into account all shortcomings of chloroplast data, it would be reasonable to give more weight to nuclear results, which were also in agreement with the morphological hypothesis, and accept two new subtribes: Arctiinae (comprising *Arctium-Cousinia*) and Saussureinae (*Saussurea-Jurinea*).

3.3. Phylogenetic relationships among subtribes

Undoubtedly, one of the major improvements that we obtained using Hyb-Seq data is the high phylogenetic resolution recovered along the backbone of the Cardueae tree (Fig. 4). Although in most recent Sanger-based phylogenies, 8 of the total 12 circumscribed subtribes were already monophyletic and well-defined (see Table 3; Barres et al., 2013), their interrelationships remained unresolved until now. In Barres et al. (2013), phylogenies inferred under maximum parsimony were completely undefined for the subtribal nodes. Similarly, under Bayesian inference most of the subtribe branches were anchored in polytomic structures, especially in the nuclear ITS phylogeny (Fig. 4A). In contrast, our phylogenetic reconstructions showed a significantly supported bifurcating pattern (Figs. 4C, D, and E).

Subtribes Carlininae, Cardopatiinae, and Echinopsinae were placed as the first diverging lineages within Cardueae near the root of the tribe. Unfortunately, we could not conclusively identify the sister

subtribe to the rest with confidence, due to a cyto-nuclear discordance. In the nuclear tree inferred under concatenation, Carlininae were sister to the rest of subtribes (Fig. 4C). However, individual gene trees show conflicting topologies at this part of the tree (Figs. 3 and 4E); and in the chloroplast tree the Cardopatiinae were sister to the rest of subtribes (Fig. 4D). Successive sister clades recovered with all reconstruction methods occur in the following order: Dipterocominae, Xerantheminae, Berardiinae, Staehelininae, and Onopordinae. Note that, as mentioned above, Dipterocominae appeared nested with Xerantheminae and Berardiinae with Staehelininae in the nuclear and chloroplast concatenated analyses, respectively. The grouping pattern of the rest of subtribes was slightly different on nuclear and plastid phylogenies. In nuclear trees, Carduinae was placed as sister to Arctiinae-Saussureinae-Centaureinae clade (BS = 100 and LPP = 0.99), which was subsequently bifurcated in Arctiinae-Saussureinae (BS = 100 and LPP = 1) and Centaureinae clade (Figs. 4C, E). In the chloroplast tree, Carduinae-Arctiinae-Saussureinae-Centaureinae were clustered in a supported clade (BS = 100) that internally formed an unsupported group of Carduinae-Arctiinae-Saussureinae (BS = 63) with Centaureinae as sister to the others (Fig. 4D).

Here, despite the finding that Hyb-Seq data offered bifurcating reconstructions, we detected that branch lengths recovered were relatively short (Figs. 1B, 2B). This fact may confirm the previous suspicion of a rapid subtribal lineage divergence, which would explain the traditional difficulty to resolve this part of the tree. The Phyparts analysis allowed us to go one step further and inspect the degree of gene tree discordance along nodes of the Cardueae tree. Results of Phyparts indicated that most of gene tree topologies are in conflict with the species tree in internal subtribal nodes (great red proportion; Fig. 3), in contrast to shallow nodes of genera or species relationships that showed more congruent gene tree histories (more proportion of blue; Fig. 3). We also detected that a considerable proportion of gene trees were poorly supported and uninformative (gray color in Fig. 3 for gene tree branches with BS < 30). Considering all results, possible sources of gene tree discordance at the backbone could derive from: (1) insufficient phylogenetic signal in sampled loci (e.g. short loci length, high missing positions or taxa as a result of capture failure, or poorly variable loci; Villaverde et al., 2018); (2) extinction of critical lineages; (3) rapid splitting of major subtribal clades; or (4) gene flow among ancestral lineages or ILS effect (Morales-Briones et al., 2018). It is important to note that although individual gene tree bipartitions were mainly in conflict at subtribal level, the coalescence species tree was, in general, not weakly supported (Fig. 3). Thus, this limitation is not necessary problematic to the aim of subtribal delimitation intended here.

3.4. Dating framework for tribe Cardueae

Figure 5 shows the divergence time estimates of Cardueae based on the nuclear tree obtained with HybPiper and the concatenation approach (1055 COS loci). Median age values and 95% CI are presented for each node in Supplementary Table S3. Our dating analysis revealed that tribe Cardueae could have originated between Late Eocene and Early Oligocene 34.12 Myr (29.97–40.04 95% CI). This time frame is in a middle range of the estimated age origin for the tribe in previous dated phylogenies that included several Cardueae members. On the one hand, the studies conducted by Kim et al. (2005) and Panero and Crozier (2016) reported a younger age of the tribe, 24–29 Myr and 22 Myr (10.61–33.35 95% CI), respectively. On the other hand, Barres et al. (2013) and Huang et al. (2016) found an older tribe median age with values of

40.15 Myr (35.80–44.26 95% CI) and ca. 42 Myr, respectively. Taxon sampling, calibration points, and molecular markers used could be responsible for dating differences between studies.

Notably, our estimation for the origin of tribe Cardueae matches particularly well the "Oi-1 Glaciation" that occurred at the transition between the Eocene and the Oligocene, when the Antarctic continental ice-shield was formed (Zachos et al., 2001). This event has been linked to a decline in atmospheric CO_2 over a period of 0.5 Myr (Pearson et al., 2009) and brought world-scale major climatic changes including cooler temperatures, increased seasonality, and aridity (Dupont-Nivet et al., 2007; Eldrett et al., 2009). Such global changes, in addition to produce large reorganization of paleo-biome distributions (e.g. Pound and Salzmann, 2017), would have acted as stimuli for plant evolution in other cold- or arid-adapted groups that also date from the Eocene/Oligocene boundary: tribe Gentianeae of Gentianaceae (crown node age = 34 Myr; Favre et al., 2016), tribes Aveneae, Poeae, Stipeae, and Triticeae of Poaceae (crown node ages = 32, 35, 37, and 33.8 Myr, respectively; Schubert et al., 2018), the entire family Brassicaceae (crown node age = 32.4 Myr; Hohmann et al., 2015), or genus *Hypericum* of Hypericaceae (Meseguer et al., 2015).

The subtribes originated in two different main time periods. Subtribes Carlininae, Cardopatiinae, Dipterocominae, Xerantheminae, and Berardiinae probably originated during the Oligocene (see Fig. 5 and Supplementary Table S3 for age of subtribes), whereas Echinopsinae, Staehelininae, Onopordinae, Carduinae, Arctiinae, Saussureinae, and Centaureinae originated during Middle-Late Miocene (Fig. 5 and Supplementary Table S3).

If we draw in parallel dating results obtained here and the ones obtained by Barres et al. (2013), which is the most comparable study in terms of taxon sampling, we can observe that tribe and subtribe age estimates were quite older in the latter than those we have found here (see Fig. 6). At tribal level, the different placement of the fossil *Raiguenrayun cura* (dated to 47.5 Myr, placed here to constrain subfamily Mutisioideae + subfamily Carduoideae; CP2 in Fig. 5) does not explain why we obtained a most recent tribal origin than Barres et al. (2013), who placed it in a more basal node as it is the crown node of Compositae family. Probably, the fact that we now have a more complete sampling at the basal part of the tree, including more representatives of subfamilies other than Carduoideae and more representatives of tribes within Carduoideae other than Cardueae, results in greater time distances between the origin of the family Compositae and tribe Cardueae. At subtribal level, the greatest differences between non-overlapping age ranges were detected for subtribes Onopordinae and Centaureinae, for which the median age differed among studies in 9.32 Myr and 15.38 Myr, respectively. The different number and nature of molecular markers used would be one of the main causes of the distinct dating results. We employed a set of 1055 biparentally inherited, low-copy nuclear loci that potentially provides more sequence variation, while Barres et al. (2013) used four chloroplast regions that are usually maternally inherited and have low mutation rates, which would explain the older age values reported in their estimations. In addition, it should be also noted that for those nodes constrained with minimum ages based on fossil calibration points, we obtained a better fitting between the minimum age assigned and the final dating result recovered in comparison with Barres et al. (2013); for instance, our CP5 constrained at 6 Myr resulted in 5.91 Myr (5.21–7.01 95% CI), and in Barres et al. (2013) the CP used at the stem node of *Centaurea* constrained at 5 Myr resulted in 15.54 Myr (12.10–19.96 95% CI).

In a general overview of the dated phylogenetic tree it is noteworthy that CI interval limits are wider and more overlapping among nodes at the base of the tree than those at the tips (Fig. 5). Accordingly, lower precision and resolution was obtained for estimated ages of non-Cardueae and early diverging lineages of Cardueae (e.g. Carlininae or Cardopatiinae). Certainly, the fact that the standard deviation of our most basal and secondary dated calibration point used (CP1) is higher than the other fossil-based CPs may cause larger uncertainty at the tree base. Moreover, the number of species included as Cardueae outgroups in our study is relatively poor, particularly compared with Panero and Crozier (2016) who used a wide representation of Compositae subfamilies but few representatives for each tribe. The different taxon sampling used could be the reason for differences in the chronograms obtained, resulting in a more uncertain dating of the poorly sampled nodes. As Linder et al. (2005) suggested, taxon sampling is one of the most influential factors in molecular clock dating analysis. For this reason, our node estimates at the tree base should be taken with caution, and undoubtedly, a wider sampling of other family members could improve lineages placement and dating resolution at the base of Cardueae.

3.5. Taxonomic proposal

Subtribe Carlininae Dumort., Fl. Belg.: 72. 1827. Type of the subtribe: *Carlina* L., Sp. Pl. 2: 828. 1753.

Perennial, spiny or spinescent herbs or shrubs, less often annual plants, monoecious or exceptionally dioecious (*Tugarinovia* Iljin). Leaves deeply pinnatisect, rarely entire. Capitula frequently subtended by pectinate leaf-like bracts, many-flowered, homogamous, rarely heterogamous with ligulate florets (Fig. 7A). Outer involucral bracts usually dissect; inner ones often showy, radiant and colored. Receptacle densely covered with large scales fused into a honeycombed structure, absent only in *Tugarinovia*. Corolla lobes very short, 1–3 mm. Anther filaments glabrous, appendages long and sericeous. Style very short. Achenes with parenchymatous pericarp, densely sericeous with twin hairs. Pappus of long, plumose bristles, often connate into stout scales, persistent or deciduous.

Genera included in the subtribe: *Atractylodes* DC., *Atractylis* L., *Carlina* L., *Thevenotia* DC., and *Tugarinovia* Iljin.

Geographic distribution (Fig. 8A): *Atractylis*, *Carlina*, and *Thevenotia* grow in Eurasia, North Africa, and the Irano-Turanian region; *Tugarinovia* is endemic to the deserts of Mongolia and NW China; *Atractylodes* is distributed in East Asia (China, Japan, and Korea).

Subtribe Cardopatiinae Less., Syn. Gen. Compos.: 14. 1832. Type of the subtribe: *Cardopatium* Juss., Ann. Mus. Natl. Hist. Nat. 6: 324. 1805.

Spiny perennial or annual herbs. Leaves dentate or pinnatisect. Capitula few-flowered, homogamous, densely (*Cardopatium*) or loosely (*Cousiniopsis* Nevski) clustered in corymbs (Fig. 7B). Involucral bracts with spiny pectinate-fimbriate appendages. Florets bright blue below, deep blue above; corolla lobes linear. Anther filaments glabrous. Style with very short lobes. Achenes with parenchymatous pericarp, densely sericeous. Pappus of two equal rows of short, lanceolate scales directly attached to the apical plate, persistent.

Genera included in the subtribe: *Cardopatium* Juss. and *Cousiniopsis* Nevski.

Geographic distribution (Fig. 8B): disjunct area; *Cardopatium* is endemic in the south of the Mediterranean Region, from Algeria to Greece and Turkey, and *Cousiniopsis* grows in the deserts of Middle Asia.

Subtribe Echinopsinae Cass. ex Dumort., Anal. Fam. Pl.: 1829. Type of the subtribe: *Echinops* L., Sp. Pl. 2: 814. 1753.

Perennial herbs, rarely annuals, usually very spiny. Leaves pinnatisect. Capitula one-flowered, aggregated in spherical (rarely hemispherical) secondary inflorescences (Fig. 7C). Bracts spiny. Corolla lobes often apically scarious, densely denticulate, blue or white. Anther filaments glabrous, basal appendages short, laciniate. Style with short lobes. Achenes with parenchymatous pericarp, densely sericeous. Pappus of broad, short scales directly attached to the apical plate, persistent.

This subtribe tribe includes only the genus *Echinops*.

Geographic distribution (Fig. 8C): the genus grows in the Mediterranean and the Irano-Turanian regions, with the main center of speciation in Iran. It is one of the genera that extends more to the south, reaching tropical Africa.

Subtribe Dipterocominae Garcia-Jacas & Susanna, **new subtribe**. Type of the subtribe: *Dipterocome* Fisch. & C. A. Mey., Index Seminum Horti Petropolitani 1: 26 1835.

Annual unarmed desert plants to 20 cm. Leaves entire, linear-spatulate. Capitula one-flowered and unisexual, clustered in second-order synflorescences in the axilles of leaf verticils. Bracts of the synflorescence in several rows, lanceolate, with a membranous, shortly ciliate margin. Female capitula 6–8 in the periphery of the synflorescence. Involucral bracts fused into a spiny-horned structure acrescent with the achene. Corollas shortly bilabiate. Style with very short lobes. Achenes enclosed in the spiny-horned bracts. Pappus of few, deciduous bristles attached inside the horns. Male capitula ca. 4 in the center of the synflorescence. Bracts of the involucre fused into a basal sheath. Corolla tubulose with very short, equal lobes. Anthers shortly caudate.

Comprises only the monotypic genus *Dipterocome* Fisch. & C. A. Mey.

Geographic distribution (Fig. 8D): Deserts in the East Mediterranean and Irano-Turanian regions: Middle East (Israel, Jordan, Lebanon, and Syria), Transcaucasia (Armenia, Azerbaijan, and Georgia) and Iran.

We have interpreted the minute heads of *Dipterocome pusilla* as second-order inflorescences, formed by female and male one-flowered capitula (Fig. 7D). The bracts of the involucre of the female florets are fused into a spiny-horned structure, and the achenes are enclosed in the arcuated structure at maturity.

Subtribe Xeranthememinae Cass. ex Dumort., Anal. Fam. Pl.: 32. 1829. Type of the subtribe: *Xeranthemum* L., Sp. Pl. 2: 857. 1753

Unarmed plants, usually annual herbs (*Chardinia* Desf., *Siebera* J. Gay, *Xeranthemum* L.), rarely dwarf shrubs (*Amphoricarpos* Vis.), exceptionally rhizomatous perennials (*Shangwua* Yu J. Wang, Raab-Straube, Susanna & J. Quan Liu). Leaves always entire, velvety underneath. Capitula usually heterogamous (with the exception of *Shangwua*), central florets discoid, peripheral florets shortly bilabiate, fertile, bisexual or female. Involucral bracts usually unarmed, rarely spiny (*Siebera*). Receptacle with large scarious scales. Corolla lobes very short. Anther filaments glabrous, appendages short, laciniate. Achenes often dimorphic.

Pappus of 5–15 long tapering or subulate scales, rarely reduced to a corona in *Chardinia*, sparsely sericeous with twin hairs, persistent (Fig. 7E).

Comprises five genera: *Amphoricarpos*, *Chardinia*, *Shangwua*, *Siebera* and *Xeranthemum*.

Geographic distribution (Fig. 8E): extremely diverse and complicated. The genus *Xeranthemum* is a widespread in waste places in Eurasia, from Middle Asia to the Iberian Peninsula, more frequently in the Mediterranean region. Two other closely related genera, *Chardinia* and *Siebera*, occupy similar habitats in the Irano-Turanian region, from the Tian Shan to Turkey. *Amphoricarpos* shows a disjunct distribution in the mountains of the East Mediterranean: two species grow in the Balkans, two more in Anatolia, and one in the Caucasus. *Shangwua* is endemic to the mountains of Middle Asia and the Himalayas, from Tajikistan to China.

Subtribe Berardiinae Garcia-Jacas & Susanna, **new subtribe**. Type of the subtribe: *Berardia* Vill., Prosp. Hist. Pl. Dauphiné: 27. 1779.

Acaulescent, unarmed perennial herb (Fig. 7F). Leaves rounded, entire or denticulate, densely woolly with veins prominent beneath, green-white above. Capitula solitary, sessile, homogamous. Involucral bracts subulate, scarious, woolly, ending in a slender flattened point. Receptacle areolate. Florets yellowish or pinkish. Staminal connective very long, apiculate. Styles very long with short stigmas. Achenes oblong, glabrous, slightly sulcate. Pericarp straw-colored. Pappus of scabrid cylindric bristles retrorsely twisted, directly attached to the apical plate, persistent.

Comprises only the monotypic genus *Berardia*.

Geographic distribution (Fig. 8F): *Berardia* is a relict genus endemic to the southern Alps (SE France and NW Italy).

Subtribe Staehelininae Garcia-Jacas & Susanna, **new subtribe**. Type of the subtribe: *Staehelina* L., Sp. Pl. 2: 840. 1753.

Unarmed dwarf shrubs or subshrubs. Leaves entire or dentate-pinnatifid, linear to obovate, dark green above, white-woolly beneath. Capitula corymbose or rarely solitary, homogamous. Involucral bracts ovate to lanceolate, mucronate, sometimes minutely hirsute, unarmed. Receptacle with wide, basally connate scales. Florets whitish or pink-purple. Corolla lobes narrowly triangular, very long. Anther filaments glabrous; basal appendages very long, sericeous. Styles long, lobes straight. Achenes linear-oblong, glabrous or sericeous, with minute apical coronula. Pappus in one row of bristles basally connate into broader paleae, more or less divided apically into pinnulate fimbriae (into capillary hairs in *Staehelina dubia* L. and *S. baetica* DC.), always overtopping involucre, sometimes deciduous in a ring.

Comprises only the genus *Staehelina* (Fig. 7G).

Geographic distribution (Fig. 8G): Mediterranean region, from the Iberian Peninsula to the limits between Turkey and Iraq (*Staehelina kurdica* Merxm. & Rech. f.). Most diverse in the Eastern Mediterranean region (six species, in contrast with only two species in the Western Mediterranean).

Subtribe Onopordinae Garcia-Jacas & Susanna, **new subtribe**. Type of the subtribe: *Onopordum* L., Sp. Pl. 2: 827. 1753.

Robust biennial or perennial herbs, usually spiny-toothed, rarely weakly unarmed. Leaves lanceolate to oblong, variously dissected. Capitula cupuliform or broadly ovoid, homogamous, discoid, most often solitary or on long peduncles. Involucral bracts usually spiny, innermost sometimes scarious. Receptacle foveolate, margins of foveoles with short scales, rarely setose; receptacular bracts very often absent. Corolla lobes triangular. Anther filaments slightly papillose or glabrous. Styles very long with straight lobes. Achenes obovoid or fusiform, surface usually transversely rugulose, wrinkled or foveolate; apical plate without caruncle; insertion areole straight or lateral-abaxial. Pappus bristles in a few rows, differentiated or not, basally connate in a ring, always deciduous as a single piece (Fig. 7H).

Comprises seven genera: *Alfredia* Cass., *Ancathia* DC., *Synurus* Iljin, *Syreitschikovia* Pavlov, *Lamyropappus* Knorring & Tamamsch., *Olgaea* Iljin, *Onopordum*, and *Xanthopappus* C. Winkl.

Geographic distribution (Fig. 8H): the genus *Onopordum* is native in the Mediterranean and Irano-Turanian regions, from the Iberian Peninsula to Middle Asia, and it has been introduced as a weed in USA (California), Argentina, Chile, South Africa, and Australia. The remaining genera have narrow distributions in Middle Asia, the Tian Shan, and west China, with the exception of *Synurus*, endemic in east China, Japan, and Korea.

Subtribe Carduinae Dumort., Fl. Belg.: 73. 1827. Type of the subtribe: *Carduus* L., Sp. Pl. 2: 820. 1753.

Perennial, biennial or annual spiny herbs or subshrubs, rarely unarmed. Leaves mostly lanceolate to oblong, entire or variously dissected. Capitula globose or cupuliform, homogamous, discoid, very rarely peripheral florets sterile and radiant (*Galactites* Moench). Involucral bracts usually spiny, innermost exappendiculate or with rudimentary appendages. Receptacle densely setose. Anther filaments papillose, rarely glabrous. Styles very long with straight lobes. Achenes obovoid-fusiform, laterally compressed, with insertion areole straight, basal or lateral-abaxial; apical plate very often slanted, inclined adaxially, usually with caruncle, missing only in *Notobasis*. Pappus inserted on a parenchymatous ring in the apical plate, with bristles in many weakly differentiated or undifferentiated rows, deciduous as a single piece (Fig. 7I).

Comprises nine genera: *Carduus*, *Cirsium* Mill., *Cynara* L., *Galactites*, *Lamyropsis* (Kharadze) Dittrich, *Notobasis* Cass., *Ptilostemon* Cass., *Silybum* Vaill., and *Tyrimnus* Cass.

Geographic distribution (Fig. 8I): Widespread in all Eurasia, especially in the Mediterranean region, extending southwards to Tropical Africa, introduced elsewhere as very noxious weeds. One genus, *Cirsium*, native in America, from Canada to Guatemala.

Subtribe Saussureinae Garcia-Jacas & Susanna, **new subtribe**. Type of the subtribe: *Saussurea* DC., Ann. Mus. Natl. Hist. Nat. 16: 156. 1810.

Unarmed perennial herbs or subshrubs, very rarely annual herbs. Leaves entire or pinnatisect, often silver-white below and glabrous above, sometimes hirsute-scabrid. Capitula cylindrical or globose, often paniculate, homogamous, discoid. Involucral bracts with short appendages, unarmed. Receptacle densely setose. Anther filaments glabrous, rarely papillose. Styles long, with reflexed lobes. Achenes cylindrical, slightly obconical or obpyramidal, indistinctly ribbed to costate, surface smooth or transversally rugose, very rarely with spines or scales, with or without a crown; apical plate with a persisting style base, without caruncle. Pappus of very long (overtopping involucral bracts), showy, scabrid or plumose, strongly

differentiated bristles, the inner ones basally connate in a ring, persistent or detachable as a single piece, outer bristles shorter and freely deciduous (Fig. 7J).

Comprises four genera: *Dolomiaea* DC., *Jurinea* Cass., Polytaxis C. Winkl., and *Saussurea* DC.

Geographic distribution (Fig. 8J): Native in Eurasia, *Jurinea* grows especially in the eastern Mediterranean and West Asia with three species reaching the mountains of the Iberian Peninsula and Morocco; *Saussurea* has its most important center of speciation in the Himalayas and the Hengduan mountains (more than 300 species); two species in Western Europe (Alps and Pyrenees), and six species in North America. *Dolomiaea* is circumscribed to the Himalayas and adjacent mountains. *Polytaxis* is limited to the Tian Shan in Middle Asia.

Subtribe Arctiinae Garcia-Jacas & Susanna, **new subtribe**. Type of the subtribe: *Arctium* L., Sp. Pl. 2: 816. 1753.

Biennial or perennial herbs, polycarpic or rarely monocarpic, exceptionally annuals, very spiny, less often unarmed. Leaves entire, lyrate or pinnatisect, usually with arachnoid pubescence. Capitula homogamous, cylindrical or globose, discoid. Outer and middle involucral bracts spiny or hooked, inner bracts ended in a scarious, unarmed or spiny appendage. Receptacle with long, twisted fimbrillae. Anther filaments slightly papillose or glabrous. Styles with short, straight lobes. Achenes cylindric to narrowly obovoid, usually laterally compressed or 4-costate, rarely shortly winged, usually longitudinally inconspicuously ribbed and stripped, apex glabrous or coronulate; apical plate truncate, without caruncle; insertion areole lateral-adaxial. Pappus in three rows of freely deciduous undifferentiated bristles, rarely reduced (Fig. 7K).

Comprises two genera: *Arctium* L. and *Cousinia* Cass.

Geographic distribution (Fig. 8K): some species of *Arctium* are widespread in Eurasia, but most of the species are endemic to the mountains of Middle Asia, especially the Tian Shan. *Cousinia* is confined to the Irano-Turanian floristic region, where it has radiated explosively (600 species!).

Subtribe Centaureinae Dumort., Fl. Belg.: 72. 1827. Type of the subtribe: *Centaurea* L., Sp. Pl. 2: 909. 1753.

Perennial, biennial or annual unarmed herbs, shrubs or very rarely treelets, rarely spiny. Leaves oblong to linear, entire or variously dissected. Capitula ovoid, cupuliform or cylindric, often heterogamous with sterile radiant florets, less often homogamous, discoid. Involucral bracts usually ended in a diversely scarious, fimbriate, pectinate, spiny or unarmed appendage; innermost bracts always with a scarious appendage. Receptacle densely setose. Anther filaments papillose, rarely glabrous; basal appendages long, laciniate. Style and stigma very variable. Achenes obovoid, laterally compressed, with sclerified pericarp, usually glabrous and smooth, less often hirsute, rugose, pitted, or ridged; insertion areole concave, lateral-adaxial, very rarely (*Crupina*) straight, often with an elaiosome; apical plate straight, without caruncle. Pappus inserted on a parenchymatous ring in the apical plate, biseriate, outer series formed by several rows of long, differently pinnulate bristles or rarely scales, basally connate in a ring or free, rarely deciduous as a whole of separately, often persistent, inner series of some short or longer bristles, sometimes reduced; rarely pappus single or missing by abortion (Fig. 7L).

Comprises 32 genera: *Amberboa* Vaill., *Callicephalus* C. A. Mey. *Carduncellus* Adans., *Carthamus* L., *Centaurea* L., *Centaurodendron* Johow, *Centaurothamnus* Wagenitz & Dittrich, *Cheirolophus* Cass., *Crocodilium* Vaill., *Crupina* (Pers.) DC., *Femeniasia* Susanna, *Goniocaulon* Cass., *Karvandarina* Rech. f., *Klasea* Cass., *Mantisalca* Cass., *Myopordon* Boiss., *Ochrocephala* Dittrich, *Oligochaeta* (DC.) K. Koch, *Phalacrachena* Iljin, *Phonus* Hill, *Plagiobasis* Schrenk, *Plectocephalus* D. Don, *Psephellus* Cass., *Rhaponticoides* Vaill., *Rhaponticum* Vaill., *Russowia* C. Winkl., *Schischkinia* Iljin, *Serratula* L., *Stizolophus* Cass., *Tricholepis* DC., *Volutaria* Cass., and *Zoegea* L.

Geographic distribution (Fig. 8L): mainly Mediterranean and Irano-Turanian regions, with some species of *Centaurea* extended widely in north Eurasia as far as Scotland, and the easternmost representatives reaching Middle Asia; exceptionally, *Tricholepis* has gone beyond the mountains of Middle Asia and reaches Burma (Myanmar); one species of *Rhaponticum* grows in Australia. In the south of the area, a few species grow in subtropical and tropical Africa, from Senegal to Somalia and southwards to Zimbabwe. *Plectocephalus* has a disjunct distribution in Ethiopia, U.S.A., and South America (Argentina, Brazil, and Chile). Some relict genera grow far away from the center of the distribution of the subtribe, namely *Centaurodendron* in the Juan Fernández archipelago (South Pacific Ocean), *Centaurothamnus* in Yemen, *Goniocaulon* in India, and *Ochrocephala* in Ethiopia and Sudan; all of them but *Centaurodendron* are monotypic. Some noxious weeds are widespread throughout the world, especially in regions with a Mediterranean-type climate (Australia, California, Chile, and South Africa): *Carthamus lanatus*, *Centaurea diffusa*, *C. solstitialis*, *C. stoebe*, *Rhaponticum repens*, and *Volutaria tubuliflora*.

Key to subtribes of tribe Cardueae

1. Heads always one-flowered, unisexual, clustered in the axilles of leaf verticils. Achenes enclosed in an arcuate, spiny-toothed diaspore	Dipterocominae
– Heads rarely one-flowered, always bisexual, not clustered in the axilles of leaf verticils Achenes free, not enclosed in a spiny-toothed diaspore	2
2. Capitula one-flowered, clustered in second-order spherical or hemispherical compound synflorescences	Echinopsinae
– Capitula not one-flowered, sometimes clustered in flat-topped corymbs but never forming second-order heads	3
3. Acaulescent perennials; achenes with the pappus bristles retrorsely twisted	Berardiinae
– Caulescent or acaulescent annuals or perennials; achenes with pappus of straight bristles	4
4. Leaves always entire. Achenes with pappus of 5-15 subulate, dentate or plumose scales, rarely reduced to a short corona	Xeranthemeinae
– Leaves entire or divided. Achenes with pappus of more than 15 bristles	5

5. Capitula usually subtended by pectinate-pinnatisect leaf-like bracts; corolla lobes usually 1–3 mm long; pappus of long, plumose bristles usually connate at the base into broader, robust scales		Carlininae
– Capitula subtended by entire or dentate bracts or on leafless pedicels; corolla lobes longer than 3 mm; pappus of free bristles or scales not connate at the base into broader scales		6
6. Capitula with 8–12 florets, rarely many-flowered; corollas bright blue below, deep blue above, with linear lobes; pappus of two equal rows of short lanceolate scales		Cardopatiinae
Capitula with more than 12 florets; corollas not bright blue below and deep blue above with lobes not linear; pappus of one or several rows of pinnulate or plumose long scales or bristles		7
7. Achenes rugulose, pitted or faintly velvety, never smooth, without apical caruncle or basal elaiosome		8
– Achenes usually smooth or rarely ridged, either with apical caruncle or basal elaiosome		9
8. Receptacle with long, twisted fimbrillae; achenes longitudinally stripped; pappus bristles individually deciduous		Arctiinae
– Receptacle usually naked and foveolate with short scales, rarely with straight bristles; achenes transversally rugulose; pappus bristles basally connate in a ring, detachable as a single piece		Onopordinae
9. Plants unarmed; leaves always white-woolly below; pappus very long, often overtopping involucral bracts at anthesis		10
– Plants unarmed or spiny; leaves woolly or glabrous below; pappus not overtopping involucral bracts		11
10. Shrubs or subshrubs. Corolla lobes very long, narrowly triangular; style very short, straight lobes; achenes linear-oblong, faintly sulcate; pappus in one row		Staehelininae
– Perennial herbs, rarely annuals; corolla lobes short, broadly triangular; style lobes long, reflexed; achenes cylindrical or slightly obconical or obpyramidal, smooth, echinate or rugulose; pappus in several rows		Saussureinae
11. Insertion areole of the achenes usually straight, basal or basal-abaxial. Apical plate of the achenes slanted, with caruncle (except. *Notobasis*); pappus undifferentiated, deciduous as a whole		Carduinae

– Insertion areole of the achenes usually concave, lateral-adaxial, often with an elaiosome; apical plate straight, without caruncle; pappus usually differentiated in two rows, persistent or rarely deciduous as a single piece, rarely absent by abortion — Centaureinae

Acknowledgments

Authors thank the herbaria that provided material for the study: B, BC, BCN, BEOU, DUSH, E, ERE, FRU, GA, GDA, H, LE, LP, MEM, MJG, MU, OS, SANT, TARI, TK, US, and W; and A. Pavlenko for kindly supplying the photograph of *Dipterocome pusilla*. Phylogenetic analyses were performed on the high performance cluster (HPC) of the computing facilities of University of Memphis, which is greatly acknowledged. Financial support from the Ministerio de Ciencia e Innovación (Project CGL2015-66703-P and Ph.D. grant to Sonia Herrando-Moraira) and the Catalan government ("Ajuts a grups consolidats" 2017-SGR1116) is also gratefully acknowledged.

Supplementary material and Appendices

All alignments and tree files for each dataset and appendices of divergence time analyses are deposited in Mendeley Data repository (http://dx.doi.org/10.17632/bhvv6rmyt6.1). Supplementary data associated with this article can be found, in the online version, at https://doi.org/xxxxx

References

Anderberg, A., Ghahremaninejad, F., Källersjö, M., 2007. The enigmatic genus *Dipterocome*. Compositae Newslett. 45, 23–36.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin V.M., Nikolenko, S.I., Pham, S., Prjibelski A.D., Pyshkin, A.V., Sirotkin A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pyshkin, A.V., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Barreda, V.D., Palazzesi, L., Katinas, L., Crisci, J.V., Tellería, M.C., Bremer, K., Passala, M.G., Bechis, F., Corsolini, R., 2012. An extinct Eocene taxon of the daisy family (Asteraceae): evolutionary, ecological and biogeographical implications. Ann. Bot., London, 109, 127–134. https://doi.org/10.1093/aob/mcr240

Barres, L., Sanmartín, I., Anderson, C.L., Susanna, A., Buerki, S., Galbany-Casals, M., Vilatersana, R., 2013. Reconstructing the evolution and biogeographic history of tribe Cardueae (Compositae). Am. J. Bot. 100, 867–882. https://doi.org/10.3732/ajb.1200058

Bentham, G., 1873. Compositae. In: Bentham, G., Hooker, J.D. (Eds.), *Genera plantarum*, vol. 2. Spottiswoode & Co., London, UK, pp. 163–533.

Boissier, P.E., 1849. *Diagnoses plantarum orientalium novarum*, series 1, vol. 11. Ducloux, Paris, France.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borowiec, M.L., 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ 4, e1660. https://doi.org/10.7717/peerj.1660

Bremer, K., 1987. Tribal interrelationships of the Asteraceae. Cladistics 3, 210–253. https://doi.org/10.1111/j.1096-0031.1987.tb00509.x

Bremer, K., 1994. Compositae: Cladistics and classification. Timber Press, Portland, Oregon, USA.

Bruun-Lund, S., Clement, W. L., Kjellberg, F., Rønsted, N., 2017. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). Mol. Phylogenet. Evol. 109, 93–104. https://doi.org/10.1016/j.ympev.2016.12.031

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Carlsen, M.M., Fér, T., Schmickl, R., Leong-Škorničková, J., Newman, M., Kress, W.J., 2018. Resolving the rapid plant radiation of early diverging lineages in the tropical Zingiberales: pushing the limits of genomic data. Mol. Phylogenet. Evol. 128, 55–68. https://doi.org/10.1016/j.ympev.2018.07.020

Cassini, H., 1819. [Different articles]. Dictionnaire de Sciences Naturelles. Paris. In: King, R., Dawson, H.W. (Eds., 1975), Cassini on Compositae. Oriole, New York, New York, USA.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Chevreux, B., Wetter, T., Suhai, S., 1999. Genome sequence assembly using trace signals and additional sequence information. In: Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB), vol. 99, pp. 45–56.

Dayrat, B., 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. 85, 407–415. https://doi.org/10.1111/j.1095-8312.2005.00503.x

Dittrich, M., 1977. Cynareae–systematic review. In: Heywood, V.H., Harborne, J.B., Turner, B.L. (Eds.), The biology and chemistry of the Compositae. Academic Press, London–New York–San Francisco, UK and USA, pp. 999–1015.

Dittrich, M., 1996a. Bemerkungen zur Tribuszugehörigkeit von *Berardia*. Ann. Naturhist. Mus. Wien 98 (B, suppl.), 329–342.

Dittrich, M., 1996b. Die Bedeutung morphologischer and anatomischer Achänen—Merkmale für die Systematik der Tribus Echinopeae Cass. and Carlineae Cass. Boissiera 51, 9–102.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973. https://doi.org/10.1093/molbev/mss075

993 Dupont-Nivet, G., Krijgsman, W., Langereis C.G., Abels, H.A., Dai, S., Fang, X., 2007. Tibetan Plateau
994 aridification linked to global cooling at the Eocene–Oligocene transition. Nature 445, 635–638.
995 https://doi.org/10.1038/nature05516
996 Edwards, S.V., Xi, Z., Janke, A., Faircloth, B.C., McCormack, J.E., Glenn, T.C., Zhong, B., Wu, S.,
997 Lemmon, E.M., Lemmon, A.R., Leache, A.D., Liu, L., Davis, C.C., 2016. Implementing and testing
998 the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol. Phylogenet. Evol.
999 94, 447–462. https://doi.org/10.1016/j.ympev.2015.10.027
1000 Eldrett, J.S., Greenwood, D.R., Harding, I.C., Huber, M., 2009. Increased seasonality through the Eocene
1001 to Oligocene transition in northern high latitudes. Nature 459, 969–974.
1002 https://doi.org/10.1038/nature08069
1003 Faircloth, B.C., 2015. PHYLUCE is a software package for the analysis of conserved genomic loci.
1004 Bioinformatics 32, 786–788. https://doi.org/10.1093/bioinformatics/btv646.
1005 Favre, A., Michalak, I., Chen, C.-H., Wang, J.-C., Pringle, J.S., Matuszak, S., Sun, H., Yuan, Y.-M., Struwe
1006 L., Muellner-Riehl, A.N., 2016. Out-of-Tibet: the spatio-temporal evolution of *Gentiana*
1007 (Gentianaceae). J. Biogeogr. 43, 1967–1978. https://doi.org/10.1111/jbi.12840
1008 Folk, R.A., Mandel, J.R., Freudenstein, J.V., 2015. A protocol for targeted enrichment of intron-containing
1009 sequence markers for recent radiations: A phylogenomic example from *Heuchera* (Saxifragaceae).
1010 Appl. Plant Sci. 3, 1500039. https://doi.org/10.3732/apps.1500039
1011 Folk, R.A., Mandel, J.R., Freudenstein, J.V., 2017. Ancestral gene flow and parallel organellar genome
1012 capture result in extreme phylogenomic discord in a lineage of angiosperms. Syst. Biol. 66, 320–337.
1013 https://doi.org/10.1093/sysbio/syw083
1014 Funk, V. A., Susanna, A., Stuessy, T.F., Robinson, H., 2009. Classification of Compositae. In: Funk, V.A.,
1015 Susanna, A., Stuessy, T.F., Bayer, R.J. (Eds.), 2009. Systematics, evolution, and biogeography of
1016 Compositae. International Association for Plant Taxonomy, Vienna, pp. 171–189.
1017 García, N., Folk, R.A., Meerow, A.W., Chamala, S., Gitzendanner, M.A., de Oliveira, R.S., Soltis, D.E.,
1018 Soltis, P.S., 2017. Deep reticulation and incomplete lineage sorting obscure the diploid phylogeny of
1019 rain-lilies and allies (Amaryllidaceae tribe Hippeastreae). Mol. Phylogenet. Evol. 111, 231–247.
1020 https://doi.org/10.1016/j.ympev.2017.04.003
1021 Garcia-Jacas, N., Garnatje, T., Susanna, A., Vilatersana, R., 2002. Tribal and subtribal delimitation and
1022 phylogeny of the Cardueae (Asteraceae): a combined nuclear and chloroplast DNA analysis. Mol.
1023 Phylogenet. Evol. 22, 51– 64. https://doi.org/10.1006/mpev.2001.1038
1024 Gardner, A.G., Sessa, E.B., Michener, P., Johnson, E., Shepherd, K.A., Howarth, D.G., Jabaily, R.S., 2016.
1025 Utilizing next-generation sequencing to resolve the backbone of the Core Goodeniaceae and inform
1026 future taxonomic and floral form studies. Molec. Phylogenet. Evol. 94, 605–617.
1027 https://doi.org/10.1016/j.ympev.2015.10.003
1028 Gouy, M., Guindon, S., Gascuel, O., 2009. SeaView version 4: a multiplatform graphical user interface for
1029 sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224.
1030 https://doi.org/10.1093/molbev/msp259

1031 Häffner, E., Hellwig, F.H., 1999. Phylogeny of the tribe Cardueae (Compositae) with emphasis on the
1032 subtribe Carduinae: an analysis based on ITS sequence data. Willdenowia 29, 27–39.
1033 https://doi.org/10.3372/wi.29.2902
1034 Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from
1035 genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucleic Acids
1036 Res. 41, e129. https://doi.org/10.1093/nar/gkt371
1037 Harris, R.S., 2007. Improved pairwise alignment of genomic DNA. PhD Thesis, Pennsylvania State
1038 University, Pennsylvania.
1039 Harrison, N., Kidner, C.A., 2011. Next-generation sequencing and systematics: what can a billion base
1040 pairs of DNA sequence data do for you? Taxon 60, 1552–1566.
1041 Heckenhauer, J., Samuel, R., Ashton, P.S., Salim, K.A., Paun, O., 2018. Phylogenomics resolves
1042 evolutionary relationships and provides insights into floral evolution in the tribe Shoreeae
1043 (Dipterocarpaceae). Mol. Phylogenet. Evol. 127, 1–13. https://doi.org/10.1016/j.ympev.2018.05.010
1044 Heilbuth, J.C., 2000. Lower species richness in dioecious clades. Am. Nat. 156, 221–241.
1045 https://doi.org/10.1086/303389
1046 Herrando-Moraira, S., The Cardueae Radiations Group (Calleja, J.A., Carnicero-Campmany, P., Fujikawa,
1047 K., Galbany-Casals, M., Garcia-Jacas, N., Im, H.-T., Kim, S.-C., Liu, J.-Q., López-Alvarado, J., López-
1048 Pujol, J., Mandel, J.R., Massó, S., Mehregan, I., Montes-Moreno, N., Pyak, E., Roquet, C., Sáez, L.,
1049 Sennikov, A., Susanna, A., Vilatersana, R.), 2018. Exploring data processing strategies in NGS target
1050 enrichment to disentangle radiations in the tribe Cardueae (Compositae). Mol. Phylogenet. Evol. 128,
1051 69–87. https://doi.org/10.1016/j.ympev.2018.07.012
1052 Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in
1053 phylogenetic analysis. Syst. Biol. 42, 182–192. https://doi.org/10.1093/sysbio/42.2.182
1054 Hoffmann, O., 1894. Compositae. In: Engler, A., Prantl, K. (Eds.), Die natürlichen Pflanzenfamilien, vol.
1055 5, part 4. Wilhelm Engelmann, Leipzig, Germany, pp. 24–333.
1056 Hohmann, N., Wolf, E.M., Lysak, M.A., Koch, M.A., 2015. A time-calibrated road map of Brassicaceae
1057 species radiation and evolutionary history. Plant Cell 27, 2770–2784.
1058 https://doi.org/10.1105/tpc.15.00482
1059 Huang, C.H., Zhang, C., Liu, M., Hu, Y., Gao, T., Qi, J., Ma, H., 2016. Multiple polyploidization events
1060 across Asteraceae with two nested events in the early history revealed by nuclear phylogenomics. Mol.
1061 Biol. Evol. 33, 2820–2835. https://doi.org/10.1093/molbev/msw157
1062 Ivanov, D., Ashraf, A.R., Utescher, T., Mosbrugger, V., Slavomirova, E., 2007. Late Miocene vegetation
1063 and climate of the Balkan region, palynology of the Beli Breg Coal Basin sediments. Geol. Carpath.
1064 58, 367–381.
1065 Jansen, R.K., Holsinger, K.E., Michaels, H.J., Palmer, J.D., 1990. Phylogenetic analysis of chloroplast
1066 DNA restriction site data at higher taxonomic levels: an example from the Asteraceae. Evolution 44,
1067 2089–2105. https://doi.org/10.1111/j.1558-5646.1990.tb04314.x
1068 Jansen, R.K., Michaels, H.J., Palmer, J.D., 1991. Phylogeny and character evolution in the Asteraceae based
1069 on chloroplast DNA restriction site mapping. Syst. Bot. 16, 98–115. https://doi.org/10.2307/2418976
1070 Jaubert, H.F., Spach, E., 1850. *Illustrationes plantarum orientalium*, vol. 3. Roret, Paris, France.

1071 Johnson, M.G., Gardner, E.M., Liu, Y., Medina, R., Goffinet, B., Shaw, A.J., Zerega, N.J.C., Wickett, N.J.,
1072 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput
1073 sequencing reads using target enrichment. Appl. Plant Sci. 4, 1600016.
1074 https://doi.org/10.3732/apps.1600016.
1075 Jones, K.E., Fér, T., Schmickl, R., Dikow, R.B., Funk, V.J., Herrando-Moraira, S., Johnston, P.R.,
1076 Siniscalchi, C.M., Susanna, A., Slovák, M., Thapa, R., Watson, L., Mandel, J.R., in preparation. An
1077 empirical assessment of a single family-wide target capture locus set at multiple evolutionary
1078 timescales within the largest family of flowering plants.
1079 Karis, P.O., Källersjö, M., Bremer, K., 1992. Phylogenetic analysis of the Cichorioideae (Asteraceae), with
1080 emphasis on the Mutisieae. Ann. Mo. Bot. Gard. 79, 416–427. https://doi.org/10.2307/2399778
1081 Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements
1082 in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010
1083 Kay, K.M., Voelckel, C., Yang, J.Y., Hufford, K.M., Kaska, D.D., Hodges, S.A., 2006. Floral characters
1084 and species diversification. In: Harder, L.D., Barrett, S.C.H. (Eds.), Ecology and evolution of flowers.
1085 Oxford University Press, Oxford, UK, pp. 311–325.
1086 Kim, K.-J., Choi, K.-S., Jansen, R.K., 2005. Two chloroplast DNA inversions originated simultaneously
1087 during the early evolution of the sunflower family (Asteraceae). Mol. Biol. Ecol. 22, 1783–1792.
1088 https://doi.org/10.1093/molbev/msi174
1089 Kim, K.J., Jansen, R.K., Wallace, R.S., Michaels, H.J., Palmer, J.D., 1992. Phylogenetic implications of
1090 rbcL sequence variation in the Asteraceae. Ann. Mo. Bot. Gard. 79, 428–445.
1091 https://doi.org/10.2307/2399779
1092 Knowles, L.L., Huang, H., Sukumaran, J., Smith, S.A., 2018. A matter of phylogenetic scale:
1093 Distinguishing incomplete lineage sorting from lateral gene transfer as the cause of gene tree discord
1094 in recent versus deep diversification histories. Am. J. Bot. 105, 376–384.
1095 https://doi.org/10.1002/ajb2.1064
1096 Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under
1097 coalescence. Syst. Biol. 56, 17–24. https://doi.org/10.1080/10635150601146041.
1098 Kück, P., Longo, G.C., 2014. FASconCAT-G: extensive functions for multiple sequence alignment
1099 preparations concerning phylogenetic studies. Front. Zool. 11, 81. https://doi.org/10.1186/s12983-014-
1100 0081-x
1101 Lagomarsino, L.P., Condamine, F.L., Antonelli, A., Mulch, A., Davis, C.C., 2016. The abiotic and biotic
1102 drivers of rapid diversification in Andean bellflowers (Campanulaceae). New Phytol. 210, 1430–1442.
1103 https://doi.org/10.1111/nph.13920
1104 Lane, A.K., Augustin, M.M., Ayyampalayam, S., Plant, A., Gleissberg, S., Di Stilio, V.S., Depamphilis
1105 C.W., Wong, G.K.-S., Kutchan, T.M., Leebens-Mack, J.H., 2018. Phylogenomic analysis of
1106 Ranunculales resolves branching events across the order. Bot. J. Linn. Soc. 187, 157–166.
1107 https://doi.org/10.1093/botlinnean/boy015
1108 Lemmon, A.R., Brown, J.M., Stanger-Hall, K., Lemmon, E.M., 2009. The effect of ambiguous data on
1109 phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Syst. Biol. 58, 130–
1110 145. https://doi.org/10.1093/sysbio/syp017

1111 Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics.
1112 Ann. Rev. Ecol. Evol. Syst. 44, 99–121. https://doi.org/10.1146/annurev-ecolsys-110512-135822
1113 Lessing, C.F., 1832. Synopsis generum Compositarum. Duncker & Humboldt, Berlin.
1114 Li, H., Durbin, R., 2009. Fast and accurate short read alignment with burrowswheeler transform.
1115 Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324
1116 Linder, H.P., Hardy, C.R., Rutschmann, F., 2005. Taxon sampling effects in molecular clock dating: an
1117 example from the African Restionaceae. Mol. Phylogenet. Evol. 35, 569–582.
1118 https://doi.org/10.1016/j.ympev.2004.12.006
1119 Liu, L., Xi, Z., Davis, C.C., 2015. Coalescent methods are robust to the simultaneous effects of long
1120 branches and incomplete lineage sorting. Mol. Biol. Evol. 32, 791–805.
1121 https://doi.org/10.1093/molbev/msu331
1122 López-Vinyallonga, S., Mehregan, I., Garcia-Jacas, N., Tscherneva, O., Susanna, A., Kadereit, J.K., 2009.
1123 Phylogeny and evolution of the *Arctium*-*Cousinia* complex (Compositae, Cardueae-Carduinae). Taxon
1124 58, 153–171.
1125 Mai, D.H., 1995. Tertiäre Vegetationsgeschichte Europas: Methoden und Ergebnisse. Gustav Fischer
1126 Verlag, Jena, Germany.
1127 Mai, D.H., 2001. Die mittelmiozänen und obermiozänen Floren aus der Meuroer und Raunoer Folge in der
1128 Lausitz, Teil II, Dicotyledonen. Palaeontogr. Abt. B 257, 35–174.
1129 Mandel, J.R., Barker, M.S., Bayer, R.J., Dikow, R.B., Gao, T., Jones, K.E., Keely, S., Kilian, N., Ma, H.,
1130 Siniscalchi, C.M., Susanna, A., Thapa, R., Watson, L., Funk, V.A., 2017. The Compositae Tree of Life
1131 in the age of phylogenomics. J. System. Evol. 55, 405–410. https://doi.org/10.1111/jse.12265.
1132 Mandel, J.R., Dikow, R.B., Funk, V.A., 2015. Using phylogenomics to resolve mega-families: an example
1133 from Compositae. J. System. Evol. 53, 391–402. https://doi.org/10.1111/jse.12167.
1134 Mandel, J.R., Dikow, R.B., Funk, V.A., Masalia, R.R., Staton, S.E., Kozik, A., Michelmore, R.W.,
1135 Rieseberg, L.H., Burke, J.M., 2014. A target enrichment method for gathering phylogenetic
1136 information from hundreds of loci: an example from the Compositae. Appl. Plant Sci. 2, 1300085.
1137 https://doi.org/10.3732/apps.1300085.
1138 Meseguer, A.S., Lobo, J.M., Ree, R., Beerling., D.J., Sanmartín, I., 2015. Integrating fossils, phylogenies,
1139 and niche models into biogeography to reveal ancient evolutionary history: the case of *Hypericum*
1140 (Hypericaceae). Syst. Biol. 64, 215–232, https://doi.org/10.1093/sysbio/syu088
1141 Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of
1142 large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE),
1143 14 Nov. 2010. New Orleans, pp. 1–8. https://doi.org/10.1109/GCE.2010.5676129.
1144 Mitchell, N., Lewis, P.O., Lemmon, E.M., Lemmon, A.R., Holsinger, K.E., 2017. Anchored
1145 phylogenomics improves the resolution of evolutionary relationships in the rapid radiation of *Protea*
1146 L. Am. J. Bot. 104, 102–115. https://doi.org/10.3732/ajb.1600227
1147 Morales-Briones, D.F., Liston, A., Tank, D.C., 2018. Phylogenomic analyses reveal a deep history of
1148 hybridization and polyploidy in the Neotropical genus *Lachemilla* (Rosaceae). New Phytol. 218, 1668–
1149 1684. https://doi.org/10.1111/nph.15099

Nicholls, J.A., Pennington, R.T., Koenen, E.J., Hughes, C.E., Hearn, J., Bunnefeld, L., Dexter, K.G., Stone, G.N., Kidner, C.A., 2015. Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (Leguminosae: Mimosoideae). Front. Plant Sci. 6, 710. https://doi.org/10.3389/fpls.2015.00710.

Ortiz, S., Bonifacino, J.M., Crisci, J.V., Funk, V.A., Hansen, H.V., Hind, D.J.N., Katinas, L., Roque, N., Sancho, G., Susanna, A., Tellería, M.C., 2009. The basal grade of the Compositae: Mutisieae (sensu Cabrera) and Carduoideae. In: Funk, V.A., Susanna, A., Stuessy, T.F., Bayer, R.J. (Eds.), Systematics, evolution, and biogeography of Compositae. International Association for Plant Taxonomy, Vienna, Austria, pp. 193–213.

Panero, J.L., Crozier, B.S., 2016. Macroevolutionary dynamics in the early diversification of Asteraceae. Mol. Phylogenet. Evol. 99, 116–132. https://doi.org/10.1016/j.ympev.2016.03.007

Panero, J.L., Funk, V.A., 2008. The value of sampling anomalous taxa in phylogenetic studies: Major clades of the Asteraceae revealed. Mol. Phylogenet. Evol. 47, 757–782.

Parks, M., Cronn, R., Liston, A., 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol. 7, 84. https://doi.org/10.1186/1741-7007-7-84

Pearson, P.N., Foster, G.L., Wade, B.S., 2009. Atmospheric carbon dioxide through the Eocene–Oligocene climate transition Nature 461, 1110–1114. https://doi.org/10.1038/nature08447

Petit, D.P., 1988. Le genre *Echinops* L. (Compositae, Cardueae). 1. Position phylétique et interprétation de l'incapitulescence. Candollea 43, 467–481.

Petit, D.P., 1997. Generic interrelationships of the Cardueae (Compositae): a cladistic analysis of morphological data. Plant Syst. Evol. 207, 173–203. https://doi.org/10.1007/BF00984388

Popescu, S.M., 2002. Repetitive changes in Early Pliocene vegetation revealed by high-resolution pollen analysis: Revised cyclostratigraphy of southwestern Romania. Rev. Palaeobot. Palynol. 120, 181–202. https://doi.org/10.1016/S0034-6667(01)00142-7

Pound M.J., Salzmann, U., 2017. Heterogeneity in global vegetation and terrestrial climate change during the late Eocene to early Oligocene transition. Sci. Rep. 7, 43386. https://doi.org/10.1038/srep43386

R Core Team, 2014. R: A language and environment for statistical computing. R Foundation for statistical Computing, Vienna.

Rambaut, A. 2016. FigTree v. 1.4.3. Department of Zoology, University of Oxford, Oxford. Available at http://tree.bio.ed.ac.uk/software/figtree/

Ranwez, V., Harispe, S., Delsuc, F., Douzery, E.J., 2011. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. PloS One 6: e22594. https://doi.org/10.1371/journal.pone.0022594

Renner, S.S., Ricklefs, R.E., 1995. Dioecy and its correlates in the flowering plants. Am. J. Bot. 82, 596–606. https://doi.org/10.1002/j.1537-2197.1995.tb11504.x

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109. https://doi.org/10.1093/oxfordjournals.molbev.a003974

1189 Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet
1190 frequencies. Mol. Biol. Evol. 33, 1654–1668. https://doi.org/10.1093/molbev/msw079
1191 Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E., Crozier, R.H., 2010. Integrative
1192 taxonomy: A multisource approach to exploring biodiversity. Ann. Rev. Entomol. 55, 421–438.
1193 https://doi.org/10.1146/annurev-ento-112408-085432
1194 Schubert, M., Marcussen, T., Meseguer, A.S., Fjellheim, S., 2018. The grass subfamily Pooideae: late
1195 Cretaceous origin and climate-driven Cenozoic diversification. bioRxiv.
1196 https://doi.org/10.1101/462440
1197 Shaw, T.I., Ruan, Z., Glenn, T.C., Liu, L., 2013. STRAW: Species TRee Analysis Web server. Nucleic
1198 Acids Res. W41, W238–W241.
1199 Small, R.L., Cronn, R.C., Wendel, J.F., 2004. Use of nuclear genes for phylogeny reconstruction in plants.
1200 Aust. Syst. Bot. 17: 145–170.
1201 Small, R.L., Ryburn, J.A., Cronn, R.C., Seelanan, T., Wendel, J.F., 1998. The tortoise and the hare:
1202 choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a
1203 recently diverged plant group. Am. J. Bot. 85, 1301–1315.
1204 Smith, S.A., Moore, M.J., Brown, J.W., Yang, Y., 2015. Analysis of phylogenomic datasets reveals
1205 conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol.
1206 15, 150. https://doi.org/10.1186/s12862-015-0423-0
1207 Smith, S.A., O'Meara, B.C., 2012. treePL: divergence time estimation using penalized likelihood for large
1208 phylogenies. Bioinformatics 28, 2689–2690. https://doi.org/10.1093/bioinformatics/bts492
1209 Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands
1210 of taxa and mixed models. Bioinformatics 22, 2688–2690.
1211 https://doi.org/10.1093/bioinformatics/btl446.
1212 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
1213 phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033
1214 Stegemann, S., Keuthe, M., Greiner, S., Bock, R., 2012. Horizontal transfer of chloroplast genomes
1215 between plant species. Proc. Natl. Acad. Sci. U.S.A., 109, 2434–2438.
1216 https://doi.org/10.1073/pnas.1114076109
1217 Stoughton, T.R., Kriebel, R., Jolles, D.D., O'Quinn, R.L., 2018. Next-generation lineage discovery: A case
1218 study of tuberous *Claytonia* L. Am. J. Bot. 105, 536–548. https://doi.org/10.1002/ajb2.1061
1219 Stubbs, R.L., Folk, R.A., Xiang, C.L., Soltis, D.E., Cellinese, N., 2018. Pseudo-parallel patterns of
1220 disjunctions in an Arctic-alpine plant lineage. Mol. Phylogenet. Evol. 123, 88–100.
1221 https://doi.org/10.1016/j.ympev.2018.02.016
1222 Susanna, A., 1987. *Femeniasia*, *novus genus Carduearum*. Collect. Bot. 17, 83–88.
1223 https://doi.org/10.3989/collectbot.1988.v17.158
1224 Susanna, A., Garcia-Jacas. N., 2007. Tribe Cardueae. In: Kadereit, J.W., Jeffrey, C. (Eds.), Flowering
1225 Plants. Eudicots. Asterales, vol. 8. In: Kubitzki, K. (Ed.), The families and genera of vascular plants.
1226 Springer Verlag, Berlin, Germany, pp. 123–146.

Susanna, A., Garcia-Jacas, N., 2009. Cardueae (Carduoideae). In: Funk, V.A., Susanna, A., Stuessy, T.F., Bayer, R.J. (Eds.), Systematics, evolution, and biogeography of Compositae. IAPT, Vienna, Austria. pp. 293–313.

Susanna, A., Garcia-Jacas, N., Hidalgo, O., Vilatersana, R. Garnatje, T., 2006. The Cardueae (Compositae) revisited: Insights from ITS, trnL-trnF, and matK nuclear and chloroplast DNA analysis. Ann. Mo. Bot. Gard. 93, 150–171. https://doi.org/10.3417/0026-6493(2006)93[150:TCCRIF]2.0.CO;2

Susanna, A., Garcia-Jacas, N., Soltis, D.E., Soltis, P.S., 1995. Phylogenetic relationships in tribe Cardueae (Asteraceae) based on ITS sequences. Am. J. Bot. 82, 1056–1068. https://doi.org/10.1002/j.1537-2197.1995.tb11571.x

Susanna, A., Garnatje, T., Garcia-Jacas, N., Vilatersana, R., 2002. On the correct subtribal placement of the genera *Syreitschikovia* and *Nikitinia* (Asteraceae, Cardueae): Carduinae or Centaureinae? Bot. J. Linn. Soc. 140, 313–319. https://doi.org/10.1046/j.1095-8339.2002.00104.x

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., Greiner, S., 2017. GeSeq–versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1), W6–W11. https://doi.org/10.1093/nar/gkx391

Vamosi, J.C., Otto, S.P., Barrett, S.C., 2003. Phylogenetic analysis of the ecological correlates of dioecy in angiosperms. J. Evol. Biol. 16, 1006–1018. https://doi.org/10.1046/j.1420-9101.2003.00559.x

Viljoen, E., Odeny, D.A., Coetzee, M.P., Berger, D.K., Rees, D.J., 2018. Application of chloroplast phylogenomics to resolve species relationships within the plant genus *Amaranthus*. J. Mol. Evol. 86, 216–239. https://doi.org/10.1007/s00239-018-9837-9

Villaverde, T., Pokorny, L., Olsson, S., Rincón-Barrado, M., Johnson, M.G., Gardner, E.M., Wickett, N.J., Molero, J., Riina, R., Sanmartín, I., 2018. Bridging the micro- and macroevolutionary levels in phylogenomics: Hyb-Seq solves relationships from populations to species and above. New Phytol. 220, 636–650. https://doi.org/10.1111/nph.15312

Wagenitz, G., 1955. Pollenmorphologie und Systematik in der Gattung *Centaurea* L. s.l. Flora 142, 213–279.

Wagenitz, G., 1958. Die Gattung *Myopordon* Boiss. (Compositae - Cynareae). Ber. Deutsch. Bot. Ges. 71, 271–277.

Wagenitz, G., 1976. Systematics and phylogeny of the Compositae (Asteraceae). Plant Syst. Evol. 125, 29–46.

Weitemier, K., Straub, S.C., Cronn, R.C., Fishbein, M., Schmickl, R., McDonnell, A., Liston, A. 2014. Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. App. Plant Sci. 2, 1400042. https://doi.org/10.3732/apps.1400042

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686–693. https://doi.org/10.1126/science.1059412

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19, 153. https://doi.org/10.1186/s12859-018-2129-y

TABLES

Table 1. Historical outline of the classification of tribe Cardueae.

Classification	Taxonomic treatment references
Tribe Echinopeae Tribe Carlineae Tribe Cardueae Subtribe Carduinae Subtribe Centaureinae	Cassini (1819), Dittrich (1977)
Tribe Cardueae Subtribe Echinopsinae Subtribe Carlininae Subtribe Carduinae Subtribe Centaureinae	Bentham (1873), Hoffmann (1894), Bremer (1994)
Tribe Echinopeae Tribe Cardueae Subtribe Carlininae Subtribe Carduinae-Centaureinae	Wagenitz (1976), Petit (1988, 1997)
Tribe Cardueae Subtribe Cardopatiinae Subtribe Carlininae Subtribe Echinopsinae Subtribe Carduinae Subtribe Centaureinae	Susanna et al. (2006), Susanna and Garcia-Jacas (2007, 2009)
Tribe Cardueae Subtribe Carlininae Subtribe Cardopatiinae Subtribe Echinopsinae Subtribe Dipterocominae Subtribe Xerantheminae Subtribe Berardiinae Subtribe Staehelininae Subtribe Onopordinae Subtribe Carduinae Subtribe Arctiinae Subtribe Saussureinae Subtribe Centaureinae	New treatment proposed here

Table 2. Some cases of genera with historical conflicts of adscription to subtribes (or even tribes) of tribe Cardueae resolved with phylogenies based on Sanger sequencing data. The tribes and treatment references in bold represent the most recent accepted taxonomic proposal and their corresponding source, respectively. The last column indicates the taxonomic treatment proposed in the present study for the traditional taxonomic conflict cases based on new subtribal delimitation. **Xeranthemum* group, including *Amphoricarpos*, *Chardinia*, *Siebera*, and *Xeranthemum*.

Genus	Taxonomic rank of conflict	Included in	References of treatments	Taxonomic treatment proposed here
Berardia	tribal	Mutisiae **Cardueae** (Carduinae)	Bremer (1994), Dittrich (1996a) **Garcia-Jacas et al. (2002)**	Cardueae (Berardiinae)
Dipterocome	tribal	Cichorieae Calenduleae **Cardueae** (Carduinae)	Boissier (1849) Jaubert & Spach (1850), Bentham (1873), Hoffmann (1894) **Susanna and Garcia-Jacas (2009)**	Cardueae (Dipterocominae)
Cardopatium	subtribal	Carlininae Echinopsidinae **Cardopatiinae**	Dittrich (1977, 1996b), Bremer (1994), Petit (1997) **Susanna et al. (2006)**	Cardopatiinae
Cousiniopsis	subtribal	Carlininae Echinopsidinae **Cardopatiinae**	Dittrich (1977, 1996b), Bremer (1994), Petit (1997) **Susanna et al. (2006)**	Cardopatiinae
Femeniasia	subtribal	Carduinae **Centaureinae**	Susanna (1987) **Susanna et al. (1995)**	Centaureinae
Myopordon	subtribal	Carduinae **Centaureinae**	Wagenitz (1958), Dittrich (1977) **Susanna et al. (2006)**	Centaureinae
Nikitinia	subtribal	Carduinae **Centaureinae**	Dittrich (1977), Bremer (1994) **Susanna et al. (2002, 2006)**	Centaureinae
Staehelina	subtribal	Carlininae **Carduinae**	Bentham (1873), Hoffmann (1894), Dittrich (1977, 1996b), Bremer (1994) Petit (1997), **Susanna et al. (2006)**	Staehelininae
Synurus	subtribal	Centaureinae **Cardueae**	Dittrich (1977), Bremer (1994) **Häffner and Hellwig (1999)**	Onopordinae
Syreitschikovia	subtribal	Centaureinae **Carduinae**	Dittrich (1977), Bremer (1994) **Susanna et al. (2002)**	Onopordinae
Xeranthemum (all the group)*	subtribal	Carlininae **Carduinae**	Dittrich (1977, 1996b), Bremer (1994), Petit (1997), **Susanna et al. (2006)**	Xerantheminae

Table 3. Overview and results from previous phylogenetic studies (performed with Sanger sequencing data) and present study (performed with NGS data) focused on tribe Cardueae with emphasis in the subtribal classification. Subtribes are divided following the classification proposed here. In the subtribe rows, the number of species sampled in the corresponding study is indicated. Statistically supported subtribes are highlighted in green, considering supported subtribes those with a monophyletic pattern and bootstrap support values ≥ 70% (in ML and MP analyses) and Bayesian posterior probabilities ≥ 0.95 (in BI analysis). In case that the study used two phylogenetic inference methods, supported subtribes are considered when one or both methods showed a significantly supported monophyly. The nuclear dataset summarized corresponds to nuclear HybPiper dataset. Abbreviations used: BI = Bayesian Inference; C = chloroplast marker; H, H = Häffner and Hellwig; ML = Maximum Likelihood; MP = Maximum Parsimony; N = nuclear marker; Nº = number; S = Susanna et al.

Systematic studies of the tribe Cardueae at subtribal level										
	S (1995)	H, H (1999)	Garcia-Jacas et al. (2002)		Susanna et al. (2006)		Barres et al. (2013)		Present study	
Nº DNA markers	1	1	1	1	1	1+2	1	4	1055	87
DNA marker type	N	N	N	C	N	N+C	N	C	N	C
Inference method	MP	MP	MP	MP	BI	MP, BI	MP, BI	MP, BI	ML	ML
Total sampling	36	32	58	42	187	108	116	124	76	67
Carlininae	1	1	**10**	**9**	11	9	**1**	**9**	**6**	**5**
Echinopsidinae	1	1	**5**	3	**7**	**5**	**7**	**7**	**3**	**2**
Cardopatiinae	-	-	2	1	**2**	**2**	**2**	**2**	**2**	**2**
Dipterocominae	-	-	-	-	-	-	1	**1**	**1**	**1**
Berardiinae	-	1	1	1	1	1	1	1	**1**	**1**
Staehelininae	-	-	1	-	**5**	**4**	**3**	**3**	**4**	**4**
Xerantheminae	-	1	**3**	**3**	**8**	**7**	**7**	**7**	**4**	**3**
Onopordiinae	-	**5**	**3**	4	**11**	**9**	**11**	**11**	**4**	**4**
Carduinae	3	14	15	11	**19**	**15**	**24**	24	**7**	**6**
Saussureinae	1	3	8	2	**25**	5	5	5	**7**	6
Arctiinae	1	**2**	5	3	28	13	10	10	**6**	**2**
Centaureinae	**29**	**4**	**5**	5	**70**	**38**	**44**	**44**	**31**	**31**

Table 4. Summary metrics of sequence extraction performance for nuclear datasets (targeting 1061 COS targets; Mandel et al., 2014) using PHYLUCE (Faircloth, 2015) and HybPiper (Johnson et al., 2016) methods, and chloroplast dataset (targeting coding regions). Parameters from 4 to 11 are calculated based on dataset "3. Nº of used loci". Parameter values were extracted with FASconCAT-G v.1.02 (Kück and Longo, 2014) and AMAS (Borowiec, 2016) programs. Abbreviations used: bp = base pairs; max = maximum; min = minimum; Nº = number; SD = standard deviation. *loci recovered for more than three species or higher than 3 bp length (see text for details).

Parameters	**Nuclear datasets**		**Chloroplast dataset**
	PHYLUCE method	**HybPiper method**	
1. Nº of species included	87	87	77
2. Nº of recovered loci (% respect COS targets)	902 (85.0)	1057 (99.6)	87
3. Nº of used loci* (% respect COS targets)	776 (73.1)	1055 (99.4)	87
4. Nº of captured loci in ≥ 90% of species (%)	13 (1.7)	729 (69.1)	85 (97.7)
5. Average of recovered loci per species (SD; min-max)	363 (59; 169–455)	924 (174; 21–1021)	84 (2; 72–86)
6. Average of species recovered per loci (SD; min-max)	41 (26; 4–82)	76 (14; 4–86)	74 (12; 4–77)
7. Mean alignment length per locus in bp (SD; min-max)	634 (360; 90–3825)	314 (184; 23–1589)	902 (1331; 89–6851)
8. Length of concatenated matrix in bp	492,549	332,260	78,531
9. Nº of variable sites in the concatenated matrix (%)	296,609 (60.2)	169,903 (51.1)	10,067 (12.8)
10. Nº of parsimony informative sites in the concatenated matrix (%)	189,716 (38.5)	123,731 (37.2)	4290 (5.5)
11. Proportion of missing data in the concatenated matrix (%)	62.6	16.6	18.2

FIGURE CAPTIONS

Fig. 1. Phylogenetic reconstruction obtained with HybPiper target extraction method and under the concatenation approach (maximum likelihood analysis performed with RAxML), showing the evolutionary relationships among 12 newly defined subtribes within tribe Cardueae. (A) cladogram, (B) phylogram. Branch labels indicate bootstrap (BS) support values, those < 70% (in red) are considered statistically unsupported.

Fig. 2. Phylogenetic reconstruction inferred from the chloroplast dataset (87 coding regions) under the concatenation approach (maximum likelihood analysis performed with RAxML), showing the evolutionary relationships among 12 newly defined subtribes within tribe Cardueae. (A) cladogram, (B) phylogram. Branch labels indicate bootstrap (BS) support values, those < 70% (in red) are considered statistically unsupported.

Fig. 3. Phylogenetic trees inferred under coalescence approach (estimated with ASTRAL) for the nuclear HybPiper, composed by 1055 COS loci. Branch labels indicate local posterior probabilities (LPP), values < 0.95 (in red) are considered statistically low supported. Results from PhyParts of conflicting and concordant gene trees are also outlined for the nuclear HybPiper dataset (see text for details). The pie charts represent the proportion of the four categories (blue, green, red, and gray) for each node. For the subtribal clades, the number above and below branches indicate the number of gene trees concordant and in conflict, respectively, with that clade respect the species tree.

Fig. 4. Summary phylogenetic trees recovered in Barres et al. (2013) and present study showing evolutionary relationships among 12 subtribes of tribe Cardueae. On the left, we show phylogenies reconstructed with nuclear sequence data (A, C, E), and on right (B, D, F), those with plastid data. Trees redrawn from Barres et al. (2013) were obtained under maximum parsimony (MP) and Bayesian inference (BI) approaches, using a nuclear dataset (A) composed by sequences of internal transcribed spacer marker (ITS) and the chloroplast dataset (B) composed by a combination of sequence data from genes *mat*K, *ndh*F, *rbc*L and intergenic spacer *trn*L-*trn*F. Phylogenetic trees inferred here derived from the nuclear HybPiper dataset composed by 1055 COS loci (C, E) and the chloroplast dataset composed by 87 coding regions (D, F), both datasets analyzed under the concatenation approach (app.; maximum likelihood [ML] analysis performed with RAxML; C, D) and coalescence approach (estimated with ASTRAL; E). Branch labels indicate: bootstrap support (BS) values in case of MP and ML analysis; posterior probabilities (pp) in BI; and local posterior probabilities (LPP) in case of under coalescence. Values with BS < 70% and pp or LPP < 0.95 (in red) are considered statistically low supported. Note that in trees of Barres et al. (2013), the low supported branches were collapsed in the redrawn version. The genus *Saussurea* and *Jurinea* belong to Saussureinae, and *Arctium* and *Cousinia* to Arctiinae. *not monophyletic subtribe.

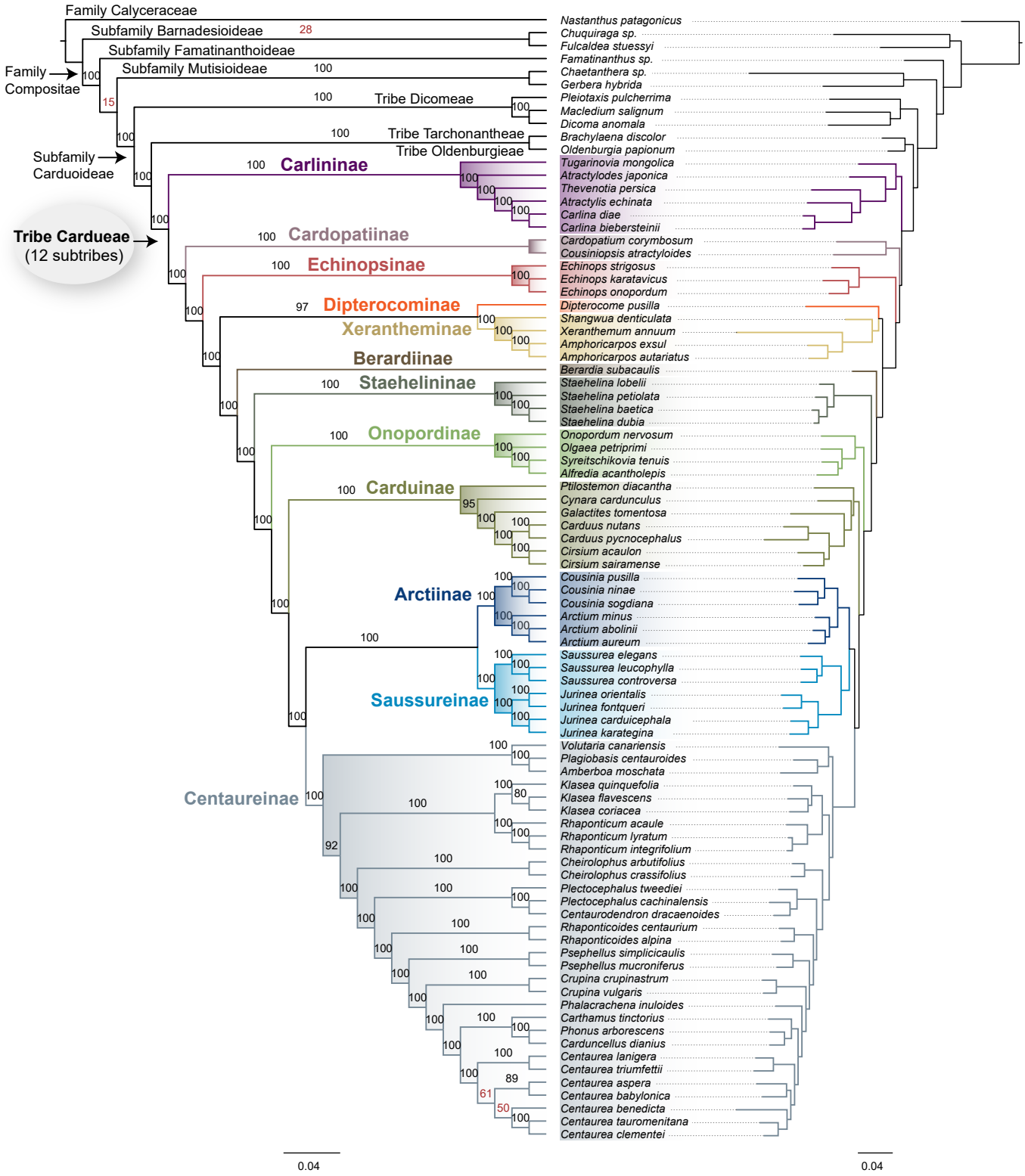
Fig. 5. Time-calibrated phylogeny of 12 newly defined subtribes within tribe Cardueae. The phylogenetic tree estimated with nuclear HybPiper dataset (1055 COS loci) and under concatenation approach (maximum likelihood analysis performed with RAxML) was used as the input tree for the penalized likelihood dating analysis using treePL. Gray bars on nodes show the 95% of confidence intervals (CI). Black circles on nodes represent the calibration points (CP) used in dating analysis. Numbers above branches are the node codes. See Supplementary Table S3 for numeric details of median estimated ages and Lower 95% CI and Upper 95% CI for each node. In the scale axis, "P" and "Q" correspond to Pliocene and Quaternary, respectively.

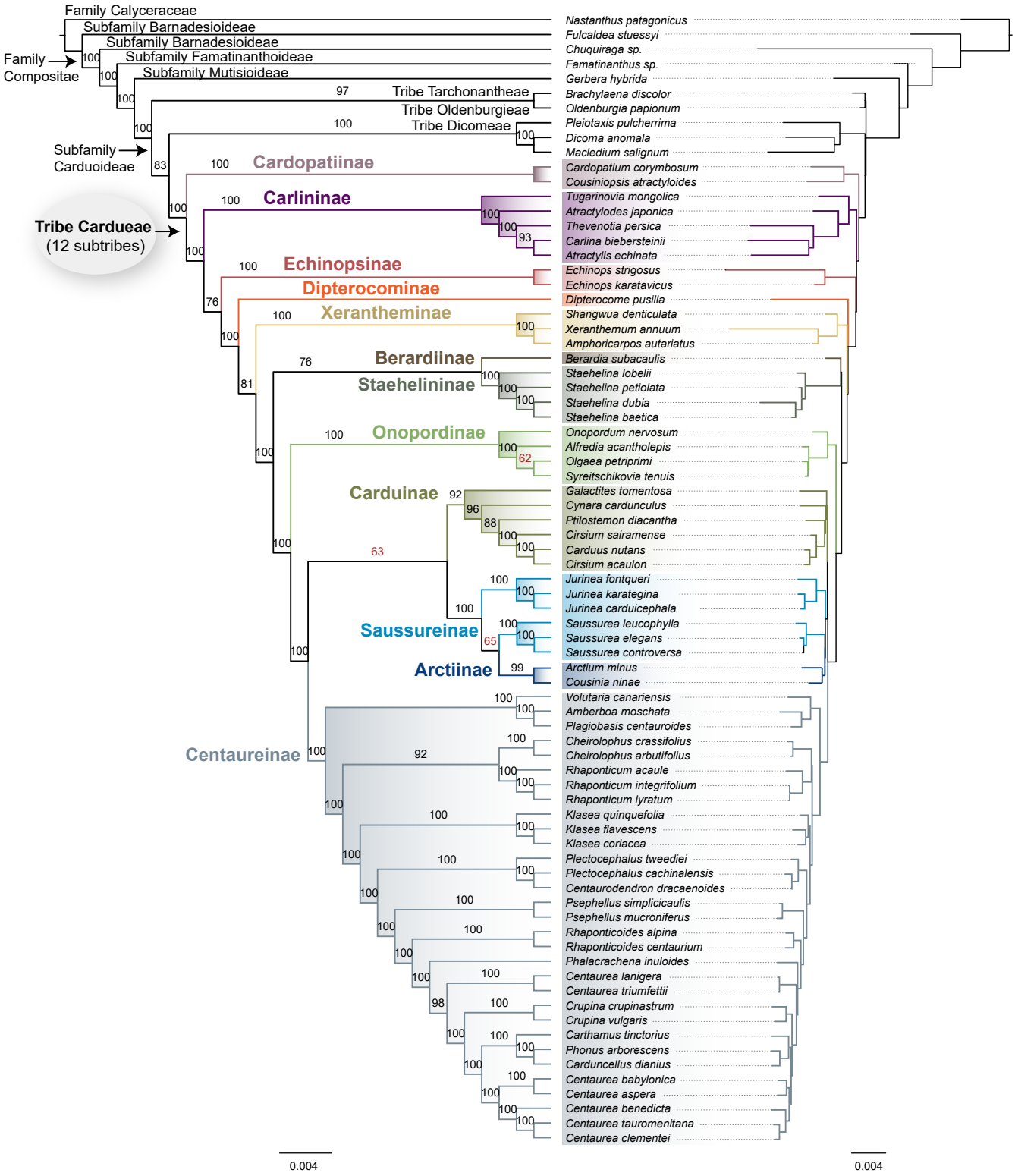
Fig. 6. Comparison of dating results of age node estimates (in Myr) for tribe Cardueae and subtribes obtained in the present study, with 1055 COS loci from nuclear genome (upper line in each subtribe), and those from Barres et al. (2013), with four chloroplast regions: *trnL-trnF*, *matK*, *ndhF*, *rbcL* (lower line in each subtribe). Colored boxes comprise the 95% of confidence interval limits and white dots represent the median age of lineages outlined. Note that subtribes Carduinae, Arctiinae, and Saussureinae are not represented due to their lack of monophyly in dated phylogeny of Barres et al. (2013).

Fig. 7. Representation of capitula diversity for 12 newly subtribes recognized here within tribe Cardueae: **(A)** subtribe Carlininae, *Atractylis humilis* L, Spain; **(B)** subtribe Cardopatiinae, *Cardopatium corymbosum* Pers., Barcelona Botanical Garden; **(C)** subtribe Echinopsinae, *Echinops viscosus* Rchb., Turkey; **(D)** subtribe Dipterocominae, *Dipterocome pusilla* Fisch. & C. A. Mey., Turkmenistan; **(E)** subtribe

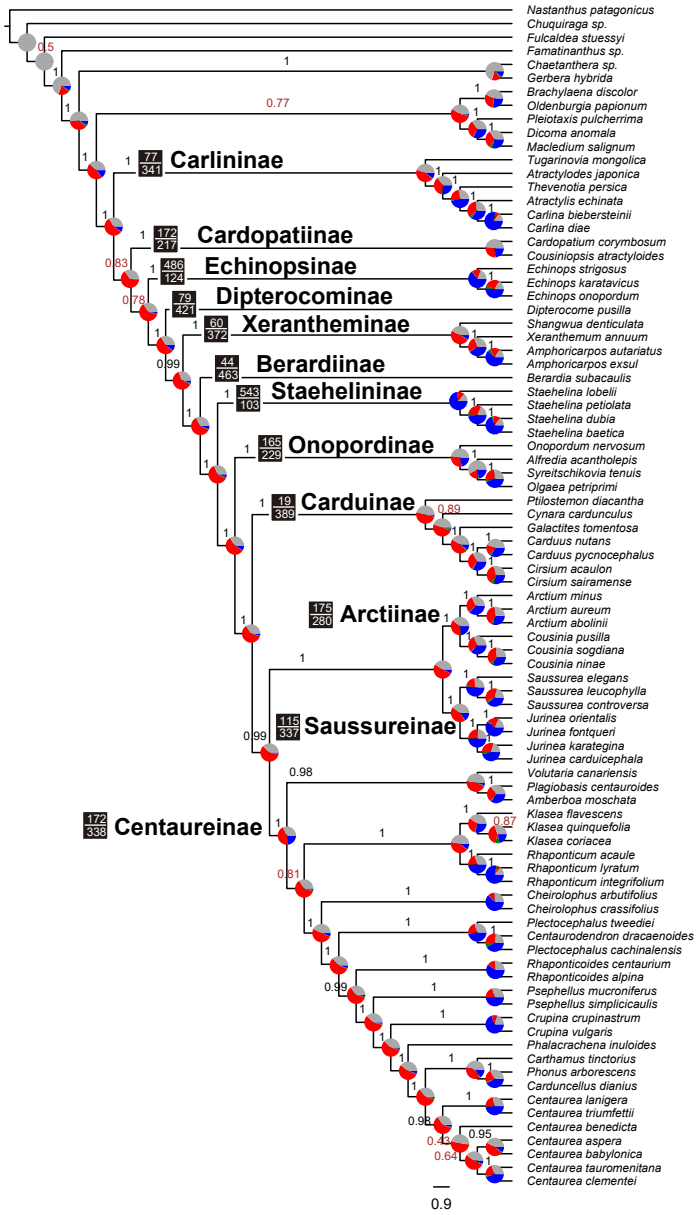
Xerantheminae, *Siebera pungens* J. Gay, Turkey; **(F)** subtribe Berardiinae, *Berardia subacaulis*, France; **(G)** subtribe Staehelininae, *Staehelina uniflosculosa* Sm., Greece; **(H)** subtribe Onopordinae, *Onopordum nervosum* Boiss., Spain; **(I)** subtribe Carduinae, *Cynara baetica* (Spreng.) Pau, Spain; **(J)** subtribe Arctiinae, *Cousinia lanata* C. Winkl., Tajikistan; **(K)** subtribe Saussureinae, *Saussurea bracteata* Decne., China; and **(L)** subtribe Centaureinae, *Centaurea cephalariifolia* Willk., Spain. (A) (B) (C) (E) (G) (H) (I) (J) (L) Photos by A. Susanna; (D) Photo by A. Pavlenko; (F) Photo by C. Roquet; (K) Photo by J. López-Pujol.

Fig. 8. Geographical representation of the main distribution ranges of 12 subtribes of tribe Cardueae defined in the present study.

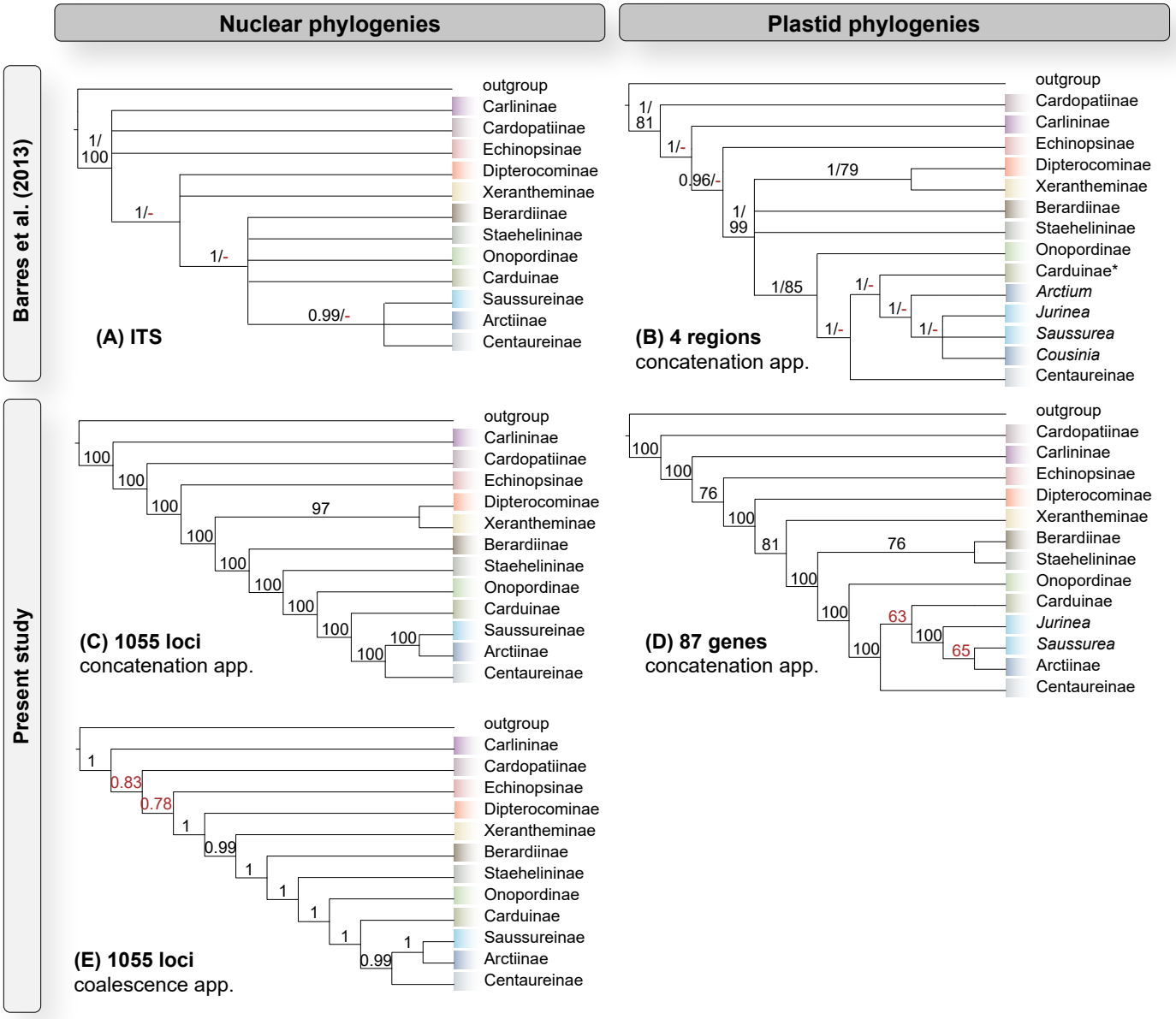

(A)
Nuclear (concatenation approach, 1055 loci)
(B)
Family Calyceraceae
Subfamily Barnadesioideae
Subfamily Famatinanthoideae
Family Compositae
Subfamily Mutisioideae
Tribe Dicomeae
Tribe Tarchonantheae
Tribe Oldenburgieae
Subfamily Carduoideae
Carlininae
Tribe Cardueae (12 subtribes)
Cardopatiinae
Echinopsinae
Dipterocominae
Xeranthemeinae
Berardiinae
Staehelininae
Onopordinae
Carduinae
Arctiinae
Saussureinae
Centaureinae
Nastanthus patagonicus
Chuquiraga sp.
Fulcaldea stuessyi
Famatinanthus sp.
Chaetanthera sp.
Gerbera hybrida
Pleiotaxis pulcherrima
Macledium salignum
Dicoma anomala
Brachylaena discolor
Oldenburgia papionum
Tugarinovia mongolica
Atractylodes japonica
Thevenotia persica
Atractylis echinata
Carlina diae
Carlina biebersteinii
Cardopatium corymbosum
Cousiniopsis atractyloides
Echinops strigosus
Echinops karatavicus
Echinops onopordum
Dipterocome pusilla
Shangwua denticulata
Xeranthemum annuum
Amphoricarpos exsul
Amphoricarpos autariatus
Berardia subacaulis
Staehelina lobelii
Staehelina petiolata
Staehelina baetica
Staehelina dubia
Onopordum nervosum
Olgaea petriprimi
Syreitschikovia tenuis
Alfredia acantholepis
Ptilostemon diacantha
Cynara cardunculus
Galactites tomentosa
Carduus nutans
Carduus pycnocephalus
Cirsium acaulon
Cirsium sairamense
Cousinia pusilla
Cousinia ninae
Cousinia sogdiana
Arctium minus
Arctium abolinii
Arctium aureum
Saussurea elegans
Saussurea leucophylla
Saussurea controversa
Jurinea orientalis
Jurinea fontqueri
Jurinea carduicephala
Jurinea karategina
Volutaria canariensis
Plagiobasis centauroides
Amberboa moschata
Klasea quinquefolia
Klasea flavescens
Klasea coriacea
Rhaponticum acaule
Rhaponticum lyratum
Rhaponticum integrifolium
Cheirolophus arbutifolius
Cheirolophus crassifolius
Plectocephalus tweediei
Plectocephalus cachinalensis
Centaurodendron dracaenoides
Rhaponticoides centaurium
Rhaponticoides alpina
Psephellus simplicicaulis
Psephellus mucroniferus
Crupina crupinastrum
Crupina vulgaris
Phalacrachena inuloides
Carthamus tinctorius
Phonus arborescens
Carduncellus dianius
Centaurea lanigera
Centaurea triumfettii
Centaurea aspera
Centaurea babylonica
Centaurea benedicta
Centaurea tauromenitana
Centaurea clementei
0.04
0.04

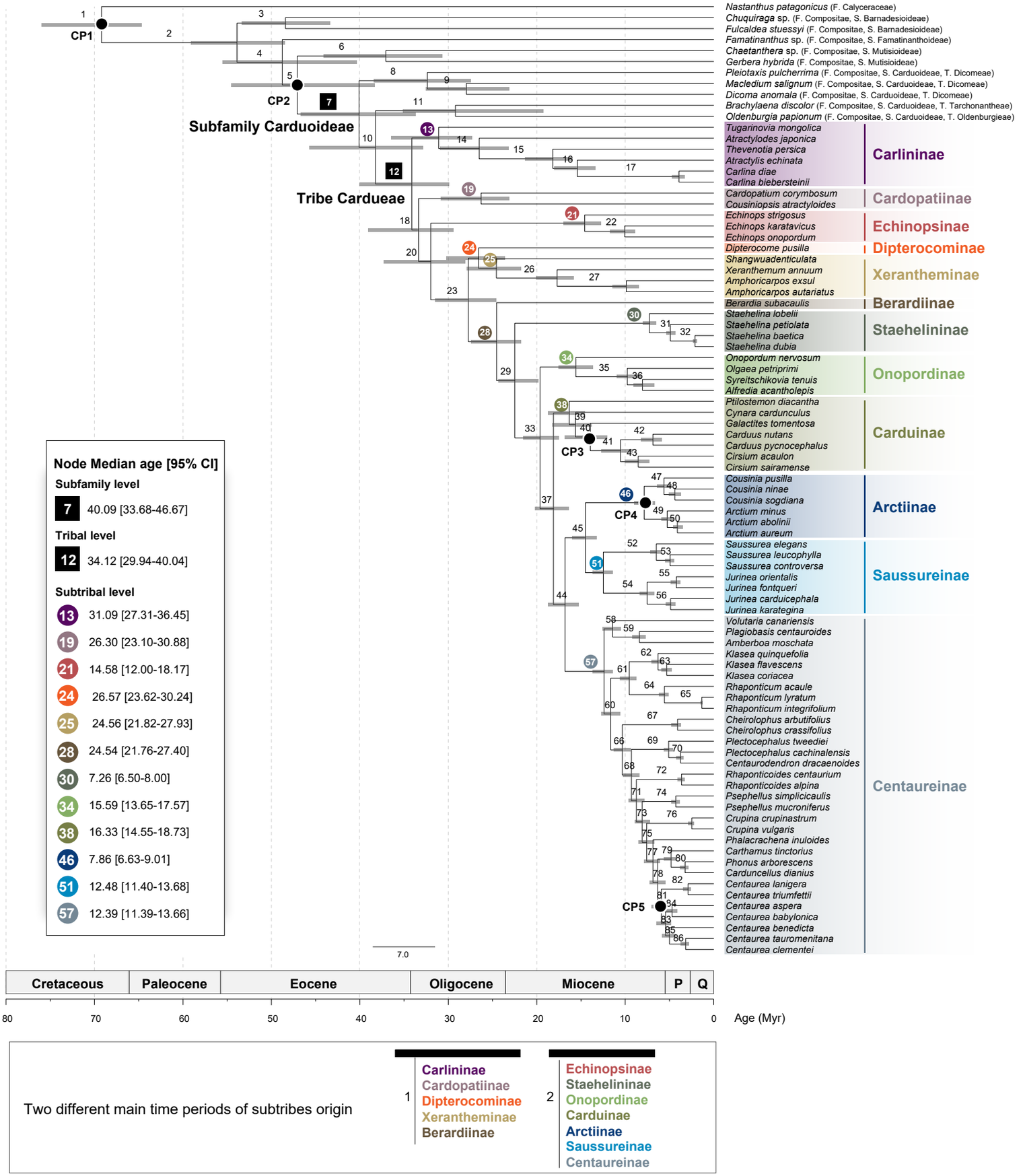

(A)
Plastid (concatenation approach, 87 genes)
(B)
Family Calyceraceae
Subfamily Barnadesioideae
Subfamily Barnadesioideae
Family Compositae
Subfamily Famatinanthoideae
Subfamily Mutisioideae
Tribe Tarchonantheae
Tribe Oldenburgieae
Tribe Dicomeae
Subfamily Carduoideae
Cardopatiinae
Carlininae
Tribe Cardueae (12 subtribes)
Echinopsinae
Dipterocominae
Xerantheminae
Berardiinae
Staehelininae
Onopordinae
Carduinae
Saussureinae
Arctiinae
Centaureinae
Nastanthus patagonicus
Fulcaldea stuessyi
Chuquiraga sp.
Famatinanthus sp.
Gerbera hybrida
Brachylaena discolor
Oldenburgia papionum
Pleiotaxis pulcherrima
Dicoma anomala
Macledium salignum
Cardopatium corymbosum
Cousiniopsis atractyloides
Tugarinovia mongolica
Atractylodes japonica
Thevenotia persica
Carlina biebersteinii
Atractylis echinata
Echinops strigosus
Echinops karatavicus
Dipterocome pusilla
Shangwua denticulata
Xeranthemum annuum
Amphoricarpos autariatus
Berardia subacaulis
Staehelina lobelii
Staehelina petiolata
Staehelina dubia
Staehelina baetica
Onopordum nervosum
Alfredia acantholepis
Olgaea petriprimi
Syreitschikovia tenuis
Galactites tomentosa
Cynara cardunculus
Ptilostemon diacantha
Cirsium sairamense
Carduus nutans
Cirsium acaulon
Jurinea fontqueri
Jurinea karategina
Jurinea carduicephala
Saussurea leucophylla
Saussurea elegans
Saussurea controversa
Arctium minus
Cousinia ninae
Volutaria canariensis
Amberboa moschata
Plagiobasis centauroides
Cheirolophus crassifolius
Cheirolophus arbutifolius
Rhaponticum acaule
Rhaponticum integrifolium
Rhaponticum lyratum
Klasea quinquefolia
Klasea flavescens
Klasea coriacea
Plectocephalus tweediei
Plectocephalus cachinalensis
Centaurodendron dracaenoides
Psephellus simplicicaulis
Psephellus mucroniferus
Rhaponticoides alpina
Rhaponticoides centaurium
Phalacrachena inuloides
Centaurea lanigera
Centaurea triumfettii
Crupina crupinastrum
Crupina vulgaris
Carthamus tinctorius
Phonus arborescens
Carduncellus dianius
Centaurea babylonica
Centaurea aspera
Centaurea benedicta
Centaurea tauromenitana
Centaurea clementei
0.004
0.004

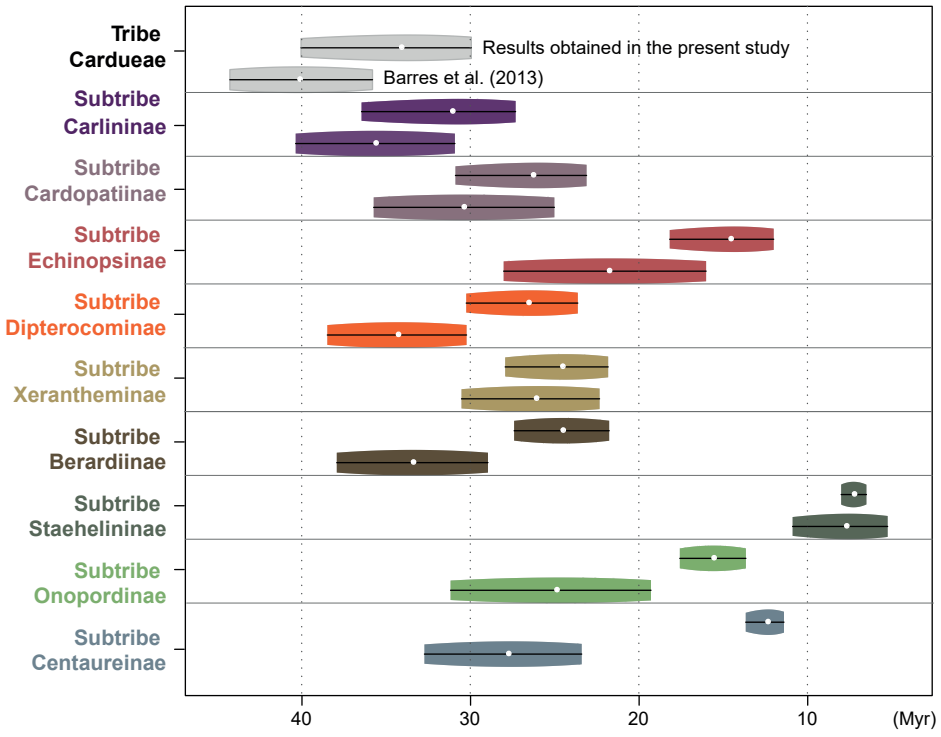
Nuclear (coalescence approach, 750 loci)

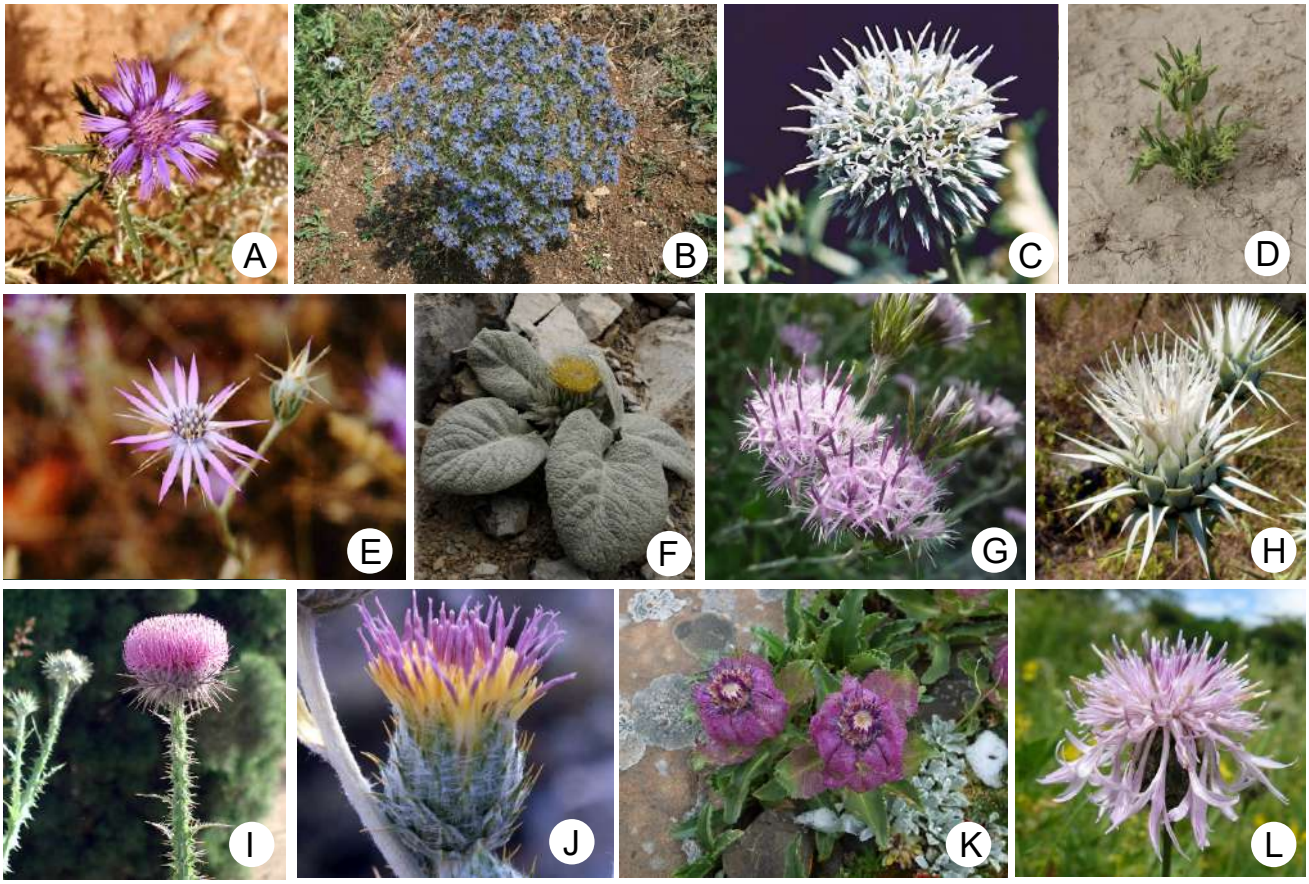


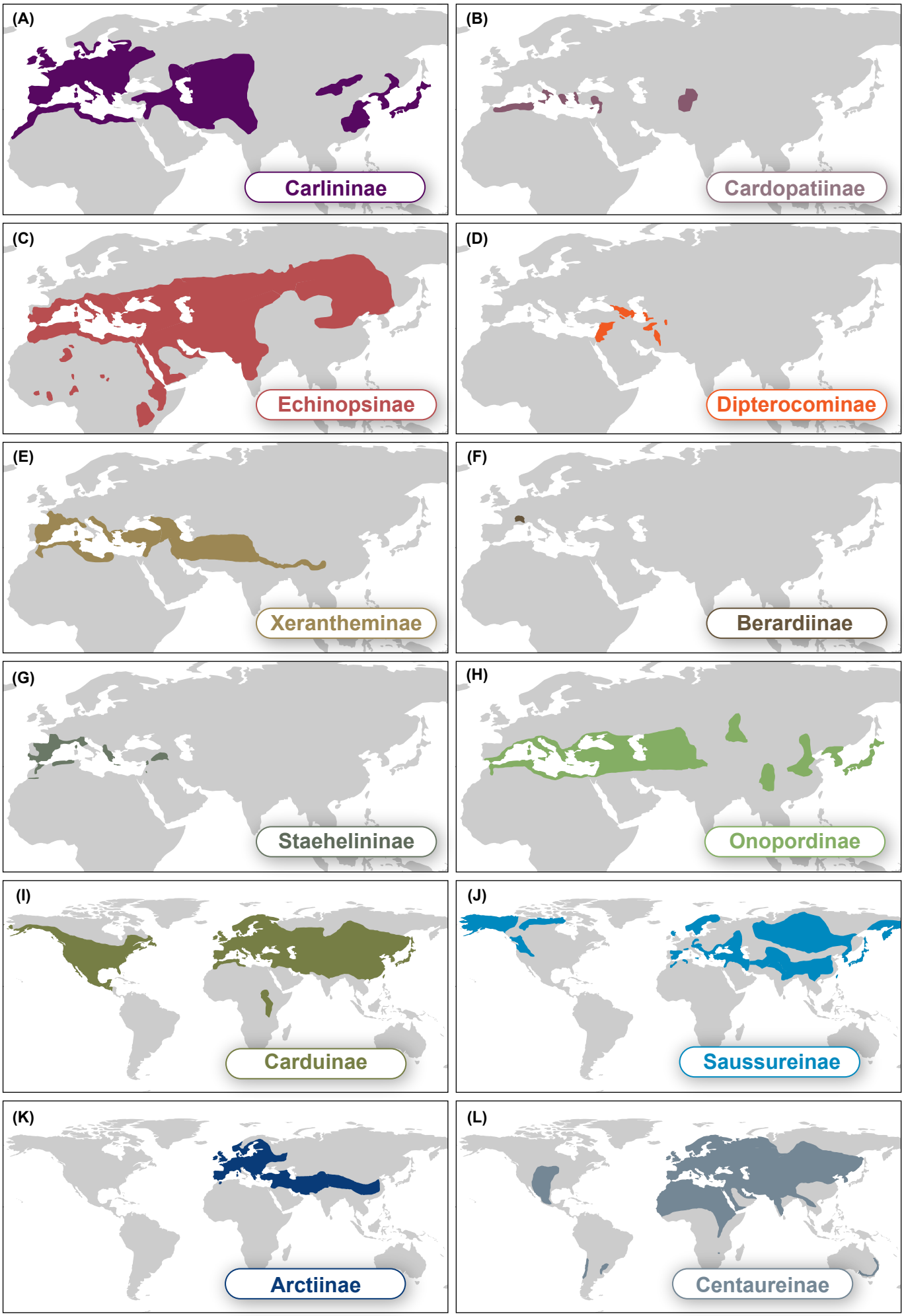
Blue: gene trees supporting that clade
Green: gene trees supporting the main alternative conflicting topology for that clade
Red: gene trees supporting the remaining conflicting topologies for that clade
Gray: gene trees supporting or confliting that clade with BS < 30%


Nuclear phylogenies
Plastid phylogenies
Barres et al. (2013)
Present study
outgroup
Carlininae
Cardopatiinae
Echinopsinae
Dipterocominae
Xerantheminae
Berardiinae
Staehelininae
Onopordinae
Carduinae
Saussureinae
Arctiinae
Centaureinae
1/ 100
1/-
1/-
0.99/-
(A) ITS
outgroup
Cardopatiinae
Carlininae
Echinopsinae
Dipterocominae
Xerantheminae
Berardiinae
Staehelininae
Onopordinae
Carduinae*
Arctium
Jurinea
Saussurea
Cousinia
Centaureinae
1/ 81
1/-
0.96/-
1/79
1/ 99
1/85
1/-
1/-
1/-
1/-
(B) 4 regions
concatenation app.
outgroup
Carlininae
Cardopatiinae
Echinopsinae
Dipterocominae
Xerantheminae
Berardiinae
Staehelininae
Onopordinae
Carduinae
Saussureinae
Arctiinae
Centaureinae
100
100
100
97
100
100
100
100
100
100
100
(C) 1055 loci
concatenation app.
outgroup
Cardopatiinae
Carlininae
Echinopsinae
Dipterocominae
Xerantheminae
Berardiinae
Staehelininae
Onopordinae
Carduinae
Jurinea
Saussurea
Arctiinae
Centaureinae
100
100
76
100
81
76
100
100
63
100
100
65
(D) 87 genes
concatenation app.
outgroup
Carlininae
Cardopatiinae
Echinopsinae
Dipterocominae
Xerantheminae
Berardiinae
Staehelininae
Onopordinae
Carduinae
Saussureinae
Arctiinae
Centaureinae
1
0.83
0.78
1
0.99
1
1
1
1
1
0.99
(E) 1055 loci
coalescence app.


Nastanthus patagonicus (F. Calyceraceae)
Chuquiraga sp. (F. Compositae, S. Barnadesioideae)
Fulcaldea stuessyi (F. Compositae, S. Barnadesioideae)
Famatinanthus sp. (F. Compositae, S. Famatinanthoideae)
Chaetanthera sp. (F. Compositae, S. Mutisioideae)
Gerbera hybrida (F. Compositae, S. Mutisioideae)
Pleiotaxis pulcherrima (F. Compositae, S. Carduoideae, T. Dicomeae)
Macledium salignum (F. Compositae, S. Carduoideae, T. Dicomeae)
Dicoma anomala (F. Compositae, S. Carduoideae, T. Dicomeae)
Brachylaena discolor (F. Compositae, S. Carduoideae, T. Tarchonantheae)
Oldenburgia papionum (F. Compositae, S. Carduoideae, T. Oldenburgieae)
Tugarinovia mongolica
Atractylodes japonica
Thevenotia persica
Atractylis echinata
Carlina diae
Carlina biebersteinii
Carlininae
Cardopatium corymbosum
Cousiniopsis atractyloides
Cardopatiinae
Echinops strigosus
Echinops karatavicus
Echinops onopordum
Echinopsinae
Dipterocome pusilla
Dipterocominae
Shangwuadenticulata
Xeranthemum annuum
Amphoricarpos exsul
Amphoricarpos autariatus
Xerantheminae
Berardia subacaulis
Berardiinae
Staehelina lobelii
Staehelina petiolata
Staehelina baetica
Staehelina dubia
Staehelininae
Onopordum nervosum
Olgaea petriprimi
Syreitschikovia tenuis
Alfredia acantholepis
Onopordinae
Ptilostemon diacantha
Cynara cardunculus
Galactites tomentosa
Carduus nutans
Carduus pycnocephalus
Cirsium acaulon
Cirsium sairamense
Carduinae
Cousinia pusilla
Cousinia ninae
Cousinia sogdiana
Arctium minus
Arctium abolinii
Arctium aureum
Arctiinae
Saussurea elegans
Saussurea leucophylla
Saussurea controversa
Jurinea orientalis
Jurinea fontqueri
Jurinea carduicephala
Jurinea karategina
Saussureinae
Volutaria canariensis
Plagiobasis centauroides
Amberboa moschata
Klasea quinquefolia
Klasea flavescens
Klasea coriacea
Rhaponticum acaule
Rhaponticum lyratum
Rhaponticum integrifolium
Cheirolophus arbutifolius
Cheirolophus crassifolius
Plectocephalus tweediei
Plectocephalus cachinalensis
Centaurodendron dracaenoides
Rhaponticoides centaurium
Rhaponticoides alpina
Psephellus simplicicaulis
Psephellus mucroniferus
Crupina crupinastrum
Crupina vulgaris
Phalacrachena inuloides
Carthamus tinctorius
Phonus arborescens
Carduncellus dianius
Centaurea lanigera
Centaurea triumfettii
Centaurea aspera
Centaurea babylonica
Centaurea benedicta
Centaurea tauromenitana
Centaurea clementei
Centaureinae
CP1
CP2
CP3
CP4
CP5
Subfamily Carduoideae
Tribe Cardueae
Node Median age [95% CI]
Subfamily level
7 40.09 [33.68-46.67]
Tribal level
12 34.12 [29.94-40.04]
Subtribal level
13 31.09 [27.31-36.45]
19 26.30 [23.10-30.88]
21 14.58 [12.00-18.17]
24 26.57 [23.62-30.24]
25 24.56 [21.82-27.93]
28 24.54 [21.76-27.40]
30 7.26 [6.50-8.00]
34 15.59 [13.65-17.57]
38 16.33 [14.55-18.73]
46 7.86 [6.63-9.01]
51 12.48 [11.40-13.68]
57 12.39 [11.39-13.66]
7.0
Cretaceous
Paleocene
Eocene
Oligocene
Miocene
P
Q
80
70
60
50
40
30
20
10
0
Age (Myr)
Two different main time periods of subtribes origin
1
Carlininae
Cardopatiinae
Dipterocominae
Xerantheminae
Berardiinae
2
Echinopsinae
Staehelininae
Onopordinae
Carduinae
Arctiinae
Saussureinae
Centaureinae


Tribe
Cardueae
Results obtained in the present study
Barres et al. (2013)
Subtribe
Carlininae
Subtribe
Cardopatiinae
Subtribe
Echinopsinae
Subtribe
Dipterocominae
Subtribe
Xeranthemeinae
Subtribe
Berardiinae
Subtribe
Staehelininae
Subtribe
Onopordinae
Subtribe
Centaureinae
40
30
20
10
(Myr)


A
B
C
D
E
F
G
H
I
J
K
L


(A)
Carlininae
(B)
Cardopatiinae
(C)
Echinopsinae
(D)
Dipterocominae
(E)
Xeranthemeinae
(F)
Berardiinae
(G)
Staehelininae
(H)
Onopordinae
(I)
Carduinae
(J)
Saussureinae
(K)
Arctiinae
(L)
Centaureinae

Supplementary Tables

Table S1. Species sampled, authority, and collection information. Groups from 1 to 12 are the Cardueae subtribes.

Taxon group	Species	Authority	Location and voucher
	Brachylaena discolor	DC.	Cultivated at the Kirstenbosch Botanical Garden.
	Chaetanthera sp.	–	Argentina, *Funk 13268* (US).
	Chuquiraga sp.	–	Chile, Arica, *Funk 13112 et al.* (US).
	Dicoma anomala	Sond.	Angola, Huambo, *S. Ortiz 853et al.* (SANT).
	Famatinanthus sp.	–	Argentina, La Rioja, *Funk 13233 et al.* (US).
	Fulcaldea stuessyi	Roque & V.A. Funk	Brazil, Bahia, *Abreu 123* (US).
	Gerbera hybrida	hort.	USA, Greenhouse grown cutting, *Mandel 105* (MEM).
	Macledium salignum	(Lawalrée) S. Ortiz	Angola, Huambo, *S. Ortiz 889 et al.* (SANT).
	Nastanthus patagonicus	Speg.	Argentina, Santa Cruz, *Bonifacino 4016 & Funk* (US).
	Oldenburgia papionum	DC.	South Africa, Western Cape, *S. Ortiz 831a* (SANT).
	Pleiotaxis pulcherrima	Klatt	Angola, Huambo, *S. Ortiz 897 et al.* (SANT).
1. Carlininae	*Atractylis echinata*	Pomel	Argelia, Naama, Djebel Antar, *Susanna s. n. et al.*, 3.VI.2010 (BC).
	Atractylodes japonica	Koidz.	Japan, Honshu, *Saito & Otomo s. n.*, 3.IX.1973 (B).
	Carlina biebersteinii	Bernh. ex Hornem.	Turkey, Konya, *Susanna 2238 et al.* (BC).
	Carlina diae	(Rech.f.) Meusel & Kästner	Cultivated at the Botanic Garden of Barcelona.
	Thevenotia persica	DC.	Iran, Tehran, *Susanna 1635B et al.* (BC).
	Tugarinovia mongolica	Iljin	Mongolia, south Gobi, *Vallès s. n. et al.*, 3.IX.2004 (BCN).
2. Cardopatiinae	*Cardopatium corymbosum*	Pers.	Greece, Macedonia: Thessaloniki, *Roché & Susanna 1951* (BC).
	Cousiniopsis atractyloides	(C.Winkl.) Nevski	Afghanistan, Kataghan, *Rechinger 33825* (E 00460517).
3. Echinopsidinae	*Echinops karatavicus*	Regel & Schmalh.	Tajikistan: Vorzov Research Station, *Susanna et al. 2458* (BC).
	Echinops onopordum	P.H. Davis	Turkey, Antalya: Tahtali Dağ, near Beycik, *Susanna 2270B et al.* (BC).
	Echinops strigosus	L.	Spain, Andalucia, *Watson 95-7A* (MU).
4. Dipterocominae	*Dipterocome pusilla*	Fisch. & C.A. Mey.	Irán, *Assadi & Maasoumi 50157* (TARI).
5. Xerantheminae	*Amphoricarpos autariatus*	Blečić & E.Mayer	Montenegro, Tara Gorge, *Janaćković s. n.*, 20.VIII.2010 (BEOU).
	Amphoricarpos exsul	O.Schwarz	Turkey, Muğla, *Susanna 2256 et al.* (BC).
	Shangwua denticulata	(DC.) Raab-Straube & Yu J. Wang	China, Xizang, Nielamu, *J.Q. Liu 07150* (BC)
	Xeranthemum annuum	L.	Turkey, Gaziantep, *Susanna 2323 et al.* (BC).
6. Berardiinae	*Berardia subacaulis*	Vill.	France, Alpes Maritimes, *C. Roquet s. n.* (Roquet pers. herb.).
7. Staehelininae	*Staehelina baetica*	DC.	Spain, Málaga, *Vilatersana 2185 et al.* (BC).
	Staehelina dubia	L.	Spain, Huesca, *Hilpold AH20084050 & López Alvarado* (BC).
	Staehelina lobelii	DC.	Turkey, Antalya, *Susanna 2272 et al.* (BC).
	Staehelina petiolata	(L.) Hilliard & B.L.Burtt	Greece, Crete, *Vitek 081418* (BC).

8. Onopordiinae	*Alfredia acantholepis*	Kar. & Kir.	Kazakhstan, Almatinskaya oblast, *Susanna 2092 et al.* (BC).
	Olgaea petriprimi	B.A.Sharipova	Tajikistan, Selandi, *Susanna 2539 et al.* (BC).
	Onopordum nervosum	Boiss.	Cultivated at the Dijon Botanical Garden.
	Syreitschikovia tenuis	(Bunge) Botsch.	Kyrgyzstan, *Seregin 2067* (FRU).
9. Carduinae	*Carduus nutans*	L.	Tajikistan, Vorzov, *Susanna 2481 et al.* (BC).
	Carduus pycnocephalus	L.	Spain, Barcelona, *Garnatje & Susanna 1827* (BC).
	Cirsium acaulon	(L.) Scop.	Tajikistan, Voru, *Susanna 2522 et al.* (BC).
	Cirsium sairamense	O. Fedtsch. & B. Fedtsch.	Tadjikistan, Maijora, *Susanna 2468 et al.* (BC).
	Cynara cardunculus	L.	USA, UW Medicinal Plant Garden., *Mandel s.n.* (GA 135).
	Galactites tomentosa	Moench	Spain, Barcelona, *Susanna s. n.*, 8.VI.2017 (BC).
	Ptilostemon diacantha	(Labill.) Greuter	Turkey, Adana, *Susanna 2313 et al.* (BC).
10. Arctiinae	*Arctium abolinii*	(Kult. ex Tscherneva) S. López et al.	Kyrgyzstan, Jalal Abad Oblast, *Lazkov s. n.* (LE).
	Arctium aureum	(C. Winkl.) Kuntze	Tajikistan, Schtut, *Susanna 2514 et al.* (BC).
	Arctium minus	(Hill) Bernh.	Spain, Barcelona, *Vilatersana 1100 & López-Vinyallonga* (BC).
	Cousinia ninae	Juz.	Kyrgyzstan, Oshskaya, *Sultanova s. n.* (LE).
	Cousinia pusilla	C. Winkl.	Tajikistan, S Tajikistan, *Botschantzev 117* (LE).
	Cousinia sogdiana	Bornm.	Uzbekistan, Karakalpakstan, *Kalibernova 5262 et al.* (LE).
11. Saussureinae	*Jurinea fontqueri*	Cuatrec.	Spain, Jaén, Mágina, *Martínez Lirola s. n.* (GDA).
	Jurinea orientalis	(Iljin) Iljin	Kyrgyzstan, Shekoftar, *Sennikov s. n.*, 13.VII.2016 (H).
	Jurinea karategina	(Lipsky) O.Fedtsch.	Tajikistan, Surjov, *Susanna 2542 et al.* (BC).
	Jurinea carduicephala	Iljin	Tajikistan: Gorno-Badakhshan, *Semakov & Dengubenko s. n.* (LE 8428).
	Saussurea controversa	DC.	Russia, Krasnoyarsk Krai, *Cazzolla Gatti 10005 et al.* (TK).
	Saussurea elegans	Ledeb.	Tajikistan, Fan mountains, *Susanna 2505 et al.* (BC).
	Saussurea leucophylla	Schrenk	Russia, Altai, *A. Pyak & E. Pyak 11073* (TK).
12. Centaureinae	*Amberboa moschata*	(L.) DC.	Armenia, *Vitek et al. 03-1539* (BC).
	Carduncellus dianius	Webb	Cultivated at the Barcelona Botanical Garden.
	Carthamus tinctorius	L.	Greenhouse grown seed, USDA PI 592391.
	Centaurea aspera	L.	Cultivated at the Barcelona Botanical Garden.
	Centaurea babylonica	M.Bieb.	Cultivated at the Barcelona Botanical Garden.
	Centaurea benedicta	(L.) L.	Greenhouse grown seed, USDA PI 311739.
	Centaurea clementei	Boiss. ex DC.	Cultivated at the Barcelona Botanical Garden.
	Centaurea lanigera	DC.	Turkey, Niğde, *Susanna 2296 et al.* (BC).
	Centaurea tauromenitana	Guss.	Cultivated at the Barcelona Botanical Garden.
	Centaurea triumfettii	All.	Turkey, Ardahan, *Susanna 2398 et al.* (BC).
	Centaurodendron dracaenoides	Johow	Chile, Juan Fernández, Masatierra, *Baeza & Peñailillo 11314* (OS).
	Cheirolophus arbutifolius	(Svent.) G.Kunkel	Cultivated at the Barcelona Botanical Garden.
	Cheirolophus crassifolius	(Bertol.) Susanna	Cultivated at the Barcelona Botanical Garden.
	Crupina crupinastrum	(Moris) Vis.	Morocco, Xauen, *Roché & Susanna 1894* (BC).

	Crupina vulgaris	Pers. ex Cass.	Spain, Palencia, *Roché & Susanna 1928* (BC).
	Klasea coriacea	(DC.) Holub	Armenia, Ararat, *Susanna 1530 et al.* (BC).
	Klasea flavescens	(L.) Holub	Spain, Almería, *Susanna 1615 et al.* (BC).
	Klasea quinquefolia	(Willd.) Greuter & Wagenitz	Iran, Azarbaijan-e Sharqi, *Susanna 1677 et al.* (BC).
	Phalacrachena inuloides	(Fisch.) Iljin	Ukraina, *Romaschenko 402 & Didukh* (BC).
	Phonus arborescens	(L.) G.López	Cultivated at the Barcelona Botanical Garden.
	Plagiobasis centauroides	Schrenk	Kazakhstan, Almaty, *Susanna 2130 et al.* (BC).
	Plectocephalus cachinalensis	(Phil.) N.Garcia & Susanna	Cultivated at the Barcelona Botanical Garden.
	Plectocephalus tweediei	(Hook. & Arn.) Garcia-Jacas & Susanna	Argentina, Chaco, *Delucchi 1970* (LP).
	Psephellus mucroniferus	(DC.) Wagenitz	Turkey, Niğde, *Susanna 2300 et al.* (BC).
	Psephellus simplicicaulis	(Boiss. & A.Huet) Wagenitz	Turkey, Erzurum, *Susanna 2386 et al.* (BC).
	Rhaponticoides alpina	(L.) M.V. Agab. & Greuter	Spain, Segovia, *Garcia-Jacas & Susanna 2798* (BC).
	Rhaponticoides centaurium	(L.) M.V. Agab. & Greuter	Italy, Apulia, *Rosati s. n.*, 22.VI.2015 (BC).
	Rhaponticum acaule	DC.	Cultivated at the Barcelona Botanical Garden.
	Rhaponticum integrifolium	C.Winkl.	Tajikistan, Vorzov, *Susanna 2478 et al.* (BC).
	Rhaponticum lyratum	Nyman	Tajikistan, Vorzov, *Susanna 2568 et al.* (BC).
	Volutaria canariensis	Wagenitz	Spain, Gran Canaria, *Kunkel 200-2* (BC).

Table S2. Species sampled and their corresponding number of raw reads, and number of recovered loci for nuclear (nrDNA) dataset (with target extraction methods: PHYLUCE from the total of 776 loci; and HybPiper from the total of 1055 loci) and chloroplast (cpDNA) dataset comprising a total of 87 genes. Groups from 1 to 12 are the Cardueae subtribes.

Taxon group	Species	Sequencing project	Nº raw reads	nrDNA		cpDNA
				Nº loci PHYLUCE	Nº loci HybPiper	Nº loci
	Brachylaena discolor	Present study	7,353,990	295	1021	84
	Chaetanthera sp.	Mandel et al. (2017)	117,364	376	445	–
	Chuquiraga sp.	Mandel et al. (2017)	39,410	192	21	72
	Dicoma anomala	Present study	4,014,387	423	952	84
	Famatinanthus sp.	Mandel et al. (2017)	1,179,957	254	991	84
	Fulcaldea stuessyi	Mandel et al. (2014)	8,999,244	358	498	82
	Gerbera hybrida	Mandel et al. (2014)	16,000,000	359	814	83
	Macledium salignum	Present study	7,513,213	455	913	84
	Nastanthus patagonicus	Mandel et al. (2017)	7,483,947	327	644	83
	Oldenburgia papionum	Present study	3,510,938	301	1017	84
	Pleiotaxis pulcherrima	Present study	7,661,670	438	971	85
1. Carlininae	*Atractylis echinata*	Present study	11,345,804	378	940	84
	Atractylis japonica	Present study	6,048,498	335	970	84
	Carlina biebersteinii	Mandel et al. (2014)	4,997,617	265	941	84
	Carlina diae	Mandel et al. (2014)	2,292,638	333	935	–
	Thevenotia persica	Present study	4,399,025	402	897	84
	Tugarinovia mongolica	Present study	3,685,187	334	981	–
2. Cardopatiinae	*Cardopatium corymbosum*	Present study	687,001	437	896	84
	Cousiniopsis atractyloides	Present study	3,918,260	419	921	84
3. Echinopsidinae	*Echinops karatavicus*	Present study	6,383,306	343	979	84
	Echinops onopordum	Present study	898,399	415	970	–
	Echinops strigosus	Mandel et al. (2014)	631,932	409	982	80
4. Dipterocominae	*Dipterocome pusilla*	Present study	4,014,451	385	973	84
5. Xeranthemeinae	*Amphoricarpos autariatus*	Present study	13,954,763	318	1010	84
	Amphoricarpos exsul	Present study	350,169	391	973	–
	Shangwua denticulata	Present study	5,295,384	315	1019	84
	Xeranthemum annuum	Present study	4,526,789	412	907	84
6. Berardiinae	*Berardia subacaulis*	Present study	4,849,918	383	1006	84
7. Staehelininae	*Staehelina baetica*	Present study	11,036,138	332	1003	84
	Staehelina dubia	Present study	3,031,747	438	996	84
	Staehelina lobelii	Present study	6,287,680	407	1000	84
	Staehelina petiolata	Present study	5,695,551	438	990	84

8. Onopordiinae	*Alfredia acantholepis*	Herrando-Moraira et al. (2018)	8,217,881	338	509	84
	Olgaea petriprimi	Herrando-Moraira et al. (2018)	5,310,933	341	1000	84
	Onopordum nervosum	Present study	2,796,798	359	989	84
	Syreitschikovia tenuis	Present study	3,390,943	359	982	85
9. Carduinae	*Carduus nutans*	Present study	3,891,005	385	959	85
	Carduus pycnocephalus	Herrando-Moraira et al. (2018)	741,845	367	666	–
	Cirsium acaulon	Present study	11,858,315	391	979	84
	Cirsium sairamense	Herrando-Moraira et al. (2018)	5,389,901	371	987	85
	Cynara cardunculus	Mandel et al. (2014)	454,885	429	796	86
	Galactites tomentosa	Present study	3,680,490	387	932	85
	Ptilostemon diacantha	Present study	5,299,804	417	998	84
10. Arctiinae	*Arctium abolinii*	Herrando-Moraira et al. (2018)	1,853,731	300	1002	–
	Arctium aureum	Herrando-Moraira et al. (2018)	4,565,685	294	1017	–
	Arctium minus	Herrando-Moraira et al. (2018)	10,007,019	349	1007	84
	Cousinia ninae	Herrando-Moraira et al. (2018)	3,280,858	303	1005	84
	Cousinia pusilla	Herrando-Moraira et al. (2018)	3,383,791	385	973	84
	Cousinia sogdiana	Herrando-Moraira et al. (2018)	3,949,050	315	999	–
11. Saussureinae	*Jurinea fontqueri*	Herrando-Moraira et al. (2018)	5,240,423	386	992	84
	Jurinea orientalis	Herrando-Moraira et al. (2018)	3,155,790	361	989	–
	Jurinea karategina	Herrando-Moraira et al. (2018)	7,713,700	393	995	85
	Jurinea carduicephala	Herrando-Moraira et al. (2018)	7,948,211	348	1013	85
	Saussurea controversa	Herrando-Moraira et al. (2018)	8,091,449	349	1011	85
	Saussurea elegans	Herrando-Moraira et al. (2018)	2,784,084	359	995	85
	Saussurea leucophylla	Herrando-Moraira et al. (2018)	5,597,695	352	1007	85
12. Centaureinae	*Amberboa moschata*	Present study	18,784,862	333	974	84
	Carduncellus dianius	Present study	5,111,933	403	943	85

	Carthamus tinctorius	Mandel et al. (2014)	10,436,332	436	857	85
	Centaurea aspera	Present study	3,215,383	365	938	85
	Centaurea babylonica	Present study	5,379,085	407	931	85
	Centaurea benedicta	Mandel et al. (2014)	644,420	423	54	85
	Centaurea clementei	Present study	9,975,987	394	947	85
	Centaurea lanigera	Present study	4,304,079	334	952	84
	Centaurea tauromenitana	Present study	4,193,748	312	940	85
	Centaurea triumfettii	Present study	6,675,594	215	976	85
	Centaurodendron dracaenoides	Present study	5,697,171	288	987	85
	Cheirolophus arbutifolius	Present study	5,547,944	413	972	84
	Cheirolophus crassifolius	Present study	5,823,163	402	982	84
	Crupina crupinastrum	Present study	3,880,017	439	949	84
	Crupina vulgaris	Present study	6,613,389	420	949	84
	Klasea coriacea	Present study	6,205,566	381	977	85
	Klasea flavescens	Present study	13,112,139	379	992	85
	Klasea quinquefolia	Present study	2,176,753	371	957	85
	Phalacrachena inuloides	Present study	5,920,329	373	985	84
	Phonus arborescens	Present study	1,975,604	404	959	85
	Plagiobasis centauroides	Present study	6,241,039	419	981	84
	Plectocephalus cachinalensis	Present study	5,898,634	169	985	85
	Plectocephalus tweediei	Present study	6,642,762	414	979	84
	Psephellus mucronifer us	Present study	2,573,295	267	963	84
	Psephellus simplicicaulis	Present study	5,370,737	229	982	85
	Rhaponticoides alpina	Present study	4,548,131	380	970	85
	Rhaponticoides centaurium	Present study	5,788,982	385	978	84
	Rhaponticum acaule	Present study	4,209,726	409	983	85
	Rhaponticum integrifolium	Present study	5,882,610	399	988	84
	Rhaponticum lyratum	Present study	7,393,929	397	975	84
	Volutaria canariensis	Present study	4,551,867	400	955	84
Average (± standard deviation)			5,420,504 (3,461,362)	364.0 (57.5)	923.4 (173.6)	84.1 (1.6)

Table S3. Median estimated ages and 95% of confidence intervals (CI) for the phylogenetic tree constructed with 1055 COS loci from nuclear HybPiper dataset by penalized likelihood method implemented in treePL software (see text for details). Node codes correspond to those represented in Figure 5.

Node code	Taxonomic unit studied	Median age (Myr)	Lower 95% CI	Upper 95% CI
1	–			
2	–	53.89	48.49	59.13
3	–	48.41	43.34	53.34
4	–	48.77	40.38	55.52
5	–	47.07	38.34	54.56
6	–	37.06	30.72	44.11
7	Subfamily Carduoideae	40.09	33.68	46.67
8	–	32.37	27.44	38.40
9	–	27.96	23.12	32.56
10	–	38.24	32.83	45.71
11	–	29.20	19.21	35.12
12	**Tribe Cardueae**	**34.12**	**29.94**	**40.04**
13	Subtribe Carlininae	**31.09**	**27.31**	**36.45**
14	–	26.51	23.20	30.97
15	–	18.19	15.93	21.30
16	–	15.41	13.37	18.01
17	–	3.92	3.30	4.71
18	–	33.36	29.40	39.05
19	Subtribe Cardopatiinae	**26.30**	**23.10**	**30.88**
20	–	31.98	28.12	37.31
21	Subtribe Echinopsinae	**14.58**	**12.00**	**18.17**
22	–	10.06	8.87	11.70
23	–	27.74	24.61	31.51
24	Subtribe Dipterocominae	**26.57**	**23.62**	**30.24**
25	Subtribe Xerantheminae	**24.56**	**21.82**	**27.93**
26	–	17.70	15.83	20.10
27	–	9.88	8.44	11.41
28	Subtribe Berardiinae	**24.54**	**21.76**	**27.40**
29	–	22.48	19.81	24.35
30	Subtribe Staehelininae	**7.26**	**6.50**	**8.00**
31	–	4.90	4.36	5.36
32	–	2.11	1.89	2.33
33	–	19.63	17.49	21.53
34	Subtribe Onopordinae	**15.59**	**13.65**	**17.57**
35	–	9.77	8.16	10.95
36	–	8.05	6.68	9.02
37	–	18.11	16.40	20.19
38	Subtribe Carduinae	**16.33**	**14.55**	**18.73**
39	–	15.62	13.78	8.22
40	–	13.95	12.05	16.88
41	–	10.53	9.18	12.72
42	–	6.86	5.88	8.25
43	–	8.53	7.29	10.08
44	–	16.80	15.23	18.72
45	–	14.50	13.23	16.02

46	Subtribe Arctiinae	**7.86**	**6.63**	**9.01**
47	–	5.62	4.77	6.42
48	–	4.39	3.70	5.04
49	–	5.25	4.46	5.89
50	–	4.10	3.47	4.57
51	Subtribe Saussureinae	**12.48**	**11.40**	**13.68**
52	–	6.48	5.89	7.20
53	–	4.93	4.46	5.51
54	–	7.52	6.75	8.38
55	–	4.22	3.75	4.85
56	–	4.87	4.33	5.38
57	Subtribe Centaureinae	**12.39**	**11.39**	**13.66**
58	–	11.42	10.52	12.55
59	–	8.39	7.71	9.21
60	–	11.63	10.60	12.71
61	–	9.55	8.70	10.58
62	–	6.29	5.65	7.01
63	–	5.31	4.79	5.92
64	–	5.57	5.10	6.18
65	–	1.34	1.23	1.50
66	–	10.33	9.37	11.29
67	–	4.08	3.67	4.78
68	–	9.29	8.39	10.26
69	–	5.09	4.60	5.58
70	–	3.74	3.40	4.19
71	–	8.74	7.82	9.64
72	–	3.63	3.30	4.12
73	–	8.09	7.22	8.96
74	–	4.28	3.85	4.83
75	–	7.58	6.76	8.52
76	–	2.48	2.26	2.90
77	–	6.85	6.09	7.85
78	–	6.32	5.46	7.27
79	–	4.81	4.25	5.59
80	–	3.23	2.82	3.80
81	–	5.91	5.21	7.01
82	–	2.89	2.58	3.43
83	–	5.44	4.82	6.43
84	–	4.70	4.13	5.50
85	–	5.00	4.36	5.83
86	–	3.17	2.79	3.75

Supplementary Figures

Figure S1. Phylogenetic reconstruction obtained with PHYLUCE target extraction method and under the concatenation approach (maximum likelihood analysis performed with RAxML), showing the evolutionary relationships among 12 newly defined subtribes within tribe Cardueae. (A) cladogram, (B) phylogram. Branch labels indicate bootstrap (BS) support values, those < 70% (in red) are considered statistically unsupported.

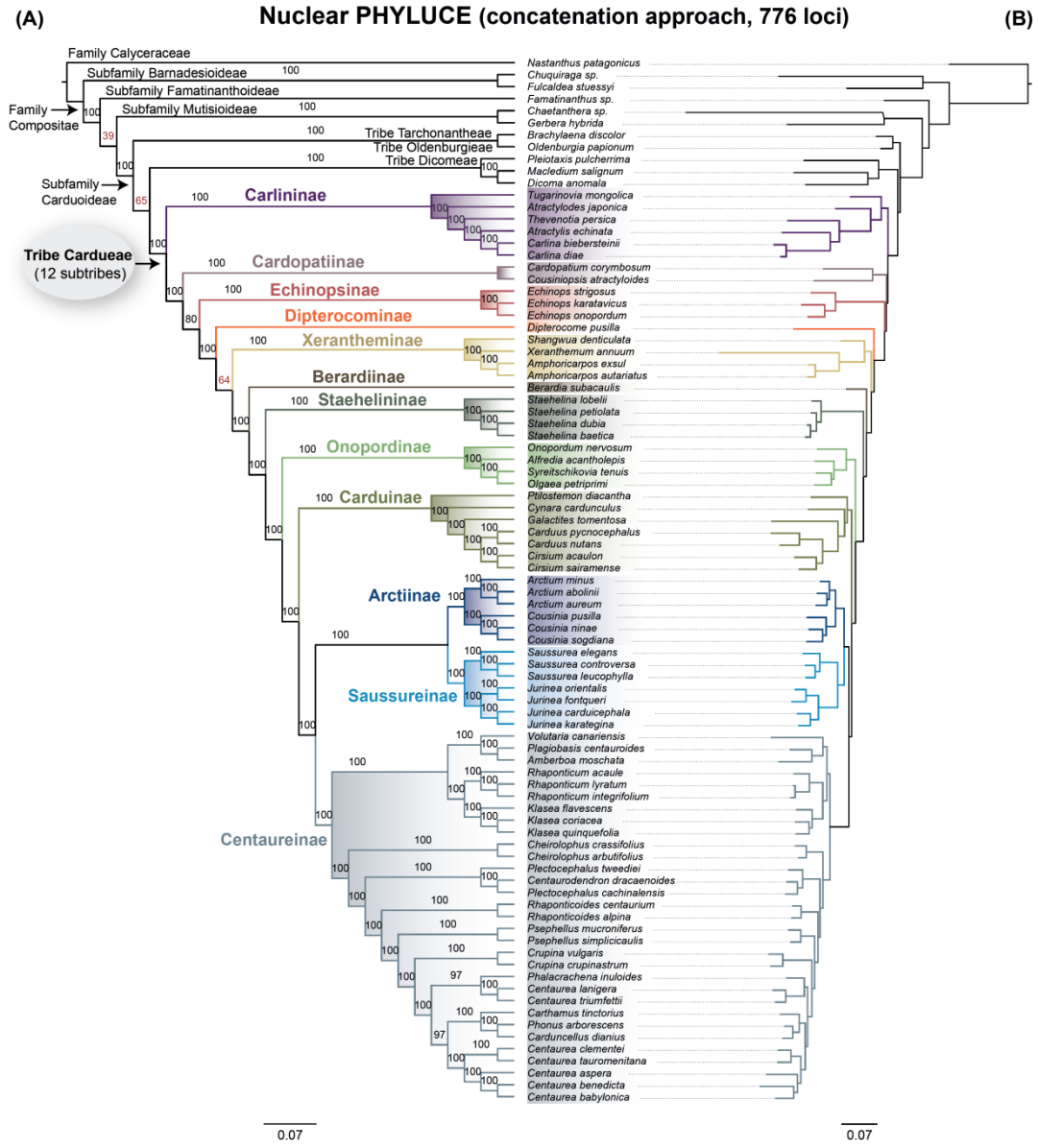
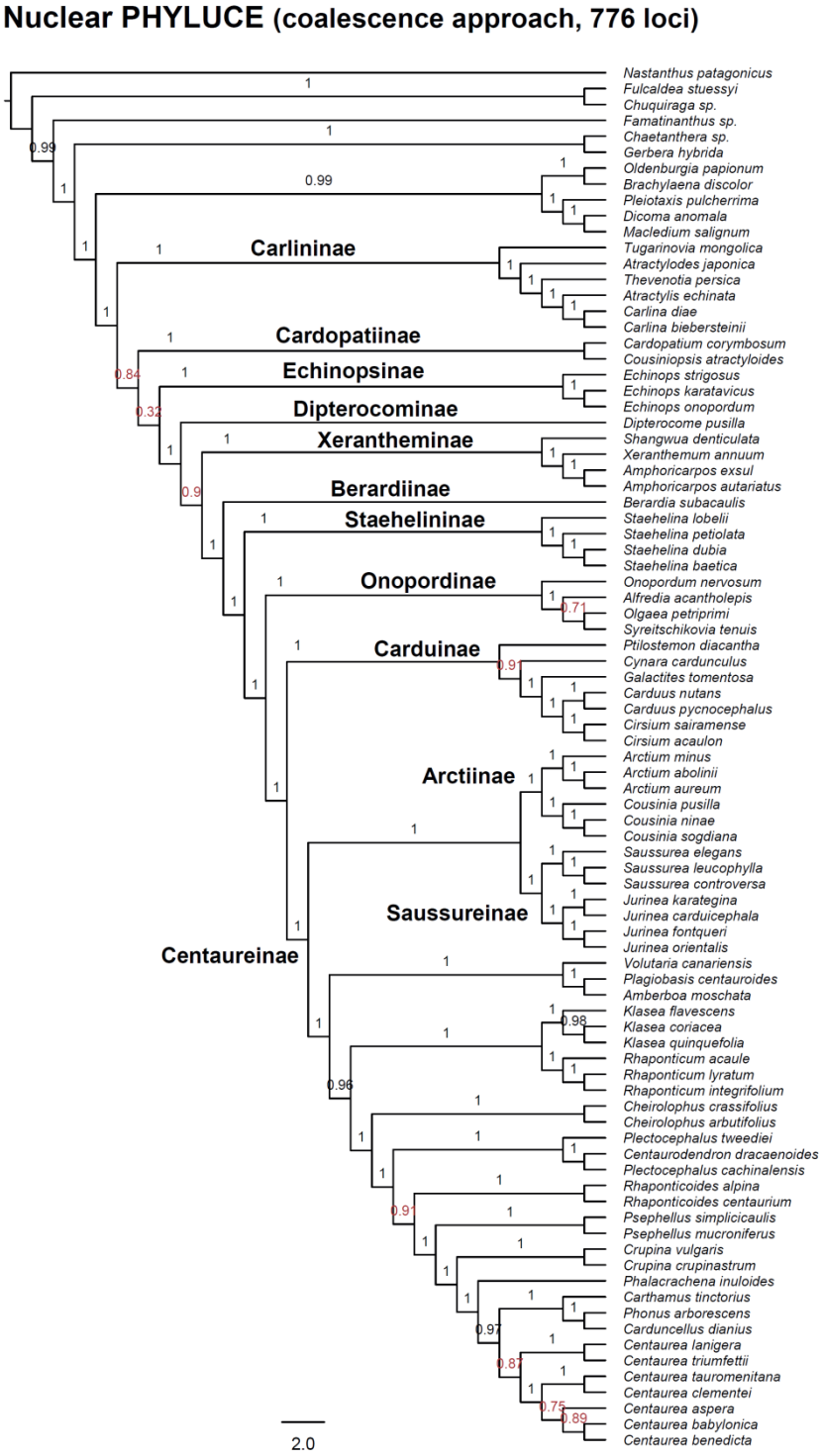



Figure S2. Phylogenetic tree inferred under coalescence approach (estimated with ASTRAL) for the nuclear PHYLUCE dataset, composed by 776 COS loci. Branch labels indicate local posterior probabilities (LPP), which values < 0.95 (in red) are considered statistically low supported.