Manuscript Details

Manuscript number	MPE_2019_268
Title	Phylogeny, origin and dispersal of <i>Saussurea</i> (Asteraceae) based on chloroplast genome data
Article type	Research Paper

Abstract

Saussurea is one of the largest genera of Asteraceae tribe Cardueae, comprising about 460 species from the Northern Hemisphere with most species distributed on the Qinghai-Tibetan Plateau (QTP) and adjacent areas. A well-supported phylogenetic framework was established based on whole chloroplast genomes of 126 taxa of *Saussurea* and 14 additional Cardueae taxa analyzed with Bayesian inference and Maximum Likelihood. Our results, however, are inconsistent with previous subgeneric classification of *Saussurea* on the whole. Some groups shown certain relationships: subgenus *Eriocoryne* was almost completely delimited, and subgenus *Theodorea*, subgenus *Saussurea* section *Laguranthera* and section *Rosulascentes* appear closely related. Molecular dating and biogeographic analysis suggest that *Saussurea* originated in Hengduan Mountains about 18.54 Mya and rapid diversification in this area and a dispersal northwards took place since then in several migration patterns. The barrier effect of QTP might be the main reason for the origin especially for the diversification of *Saussurea*. Both continuous uplift of the QTP and global cooling since mid-Miocene played important roles in the geographic expansion and diffusion; the later, especially, accelerated northward dispersal.

Keywords	<i>Saussurea</i> ; Compositae; phylogeny; evolution; biogeography; chloroplast genome
Corresponding Author	You-sheng Chen
Corresponding Author's Institution	South China Botanical Garden, Chinese Academy of Sciences
Order of Authors	Xu Liansheng, Sonia Herrando-Moraira, Alfonso Susanna, Merce Galbany Casals, You-sheng Chen
Suggested reviewers	Jun Wen, Jianquan Liu, Eckhard von Raab-Straube

Submission Files Included in this PDF

File Name [File Type]

Cover letter.doc [Cover Letter]

Highlights.docx [Highlights]

Graphical Abstract.pdf [Graphical Abstract]

text.docx [Manuscript File]

Fig. 3.pdf [Figure]

Fig. 1.tif [Figure]

Fig. 2.pdf [Figure]

Fig. 4.pdf [Figure]

Appendix Fig. 1.pdf [Supplementary Material]

Appendix Fig. 2.pdf [Supplementary Material]

Appendix Fig. 3.pdf [Supplementary Material]

Appendix Table 1.docx [Supplementary Material]

Appendix Table 2.docx [Supplementary Material]

Explanations of appendices.docx [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:
The data that has been used is confidential

1 Abstract

2 *Saussurea* is one of the largest genera of Asteraceae tribe Cardueae, comprising about
3 460 species from the Northern Hemisphere with most species distributed on the
4 Qinghai-Tibetan Plateau (QTP) and adjacent areas. A well-supported phylogenetic
5 framework was established based on whole chloroplast genomes of 126 taxa of
6 *Saussurea* and 14 additional Cardueae taxa analyzed with Bayesian inference and
7 Maximum Likelihood. Our results, however, are inconsistent with previous subgeneric
8 classification of *Saussurea* on the whole. Some groups shown certain relationships:
9 subgenus *Eriocoryne* was almost completely delimited, and subgenus *Theodorea*,
10 subgenus *Saussurea* section *Laguranthera* and section *Rosulascentes* appear closely
11 related. Molecular dating and biogeographic analysis suggest that *Saussurea* originated
12 in Hengduan Mountains about 18.54 Mya and rapid diversification in this area and a
13 dispersal northwards took place since then in several migration patterns. The barrier
14 effect of QTP might the main reason for the origin especially for the diversification of
15 *Saussurea*. Both continuous uplift of the QTP and global cooling since mid-Miocene
16 played important roles in the geographic expansion and diffusion; the later, especially,
17 accelerated northward dispersal.

18 **Key words:** *Saussurea*; Compositae; phylogeny; evolution; biogeography; chloroplast
19 genome

20 1. Introduction

21 Qinghai-Tibetan Plateau (QTP) is the earth's largest plateau, formed by several uplift
22 events after the collision of the Indian plate with Asia about 40 Mya (million years ago)
23 (Harrison et al., 1992; Spicer et al., 2003). The flora of the QTP and adjacent area has
24 a high species diversity and endemism, sustaining about 20,000 seed plant species (Wu,
25 1988; Li and Li, 1993). It harbors the highest species diversity of the North Temperate
26 zone, accounts for approximately 30% of all alpine plant species globally, and it is
27 probably the mountain system with the largest number of alpine plant species in the
28 world (Li et al. 2014; Wen et al., 2014). For these reasons, the Himalaya and the
29 Hengduan Mountains of the QTP have been listed as two of the 34 biodiversity hotspots
30 of the world (Mutke and Barthlott, 2005; Myers et al., 2000). Recent studies have
31 suggested that the uplifts of the QTP had triggered rapid diversification in both large
32 (with more than 100 species) and small plant genera (Favre et al., 2016; Mutke and
33 Barthlott, 2005; Myers et al., 2000; Qiu et al., 2011; Sun et al., 2012; Wang et al., 2009).

34 Many biogeographic works have studied the origins and migration patterns of plants in
35 the Northern Hemisphere (NH). Studies focusing on widely disjunct groups in the NH
36 have revealed at least six different biogeographic patterns (Favre et al., 2015; Wen et
37 al., 2014). One of the patterns indicates that most such genera originated on the QTP
38 and adjacent regions, and then migrated to other NH regions (including the arctic)
39 where they gave rise to derived species (Wen et al., 2014; Xu et al., 2010; Zhang et al.,
40 2009; Zhang and Fritsch, 2010). Although the results so far have illustrated complex

biogeographic connections between the QTP and other NH regions, previous studies mostly focused on woody plants (Mao et al., 2010; Yu et al., 2010), and some were based on very few molecular markers (Zhang et al., 2009; Zhang and Fritsch, 2010). In most cases, the distribution areas of the plants studied were not across all NH areas (Liu et al., 2002; Tu et al., 2010; Wen et al., 2014). Herein we investigate a highly diverse herbaceous plant group widely distributed in the NH. With its high species diversity on the QTP and the disjunct distribution in other regions of the NH, *Saussurea* DC. is an excellent model to explore the causes of high plant diversity on the QTP and the biogeographic relationships between the QTP and other NH regions.

Saussurea is one of the largest genera in the sunflower family (Asteraceae, tribe Cardueae, subtribe Saussureinae) (Herrando-Moraira et al., 2019), with about 460 species recognized up to now (Chen, 2015; Susanna and Garcia-Jacas, 2007). *Saussurea* occurs in NH, where it grows mainly in the high mountains in the Sino-Himalaya region and temperate regions of Asia, with a few species extending to North America (6 species) and Europe (9 species) (Chen, 2015; Keil, 2006; Lipschitz, 1976). The genus is especially diverse in the QTP and its adjacent regions: about 235 species (63.4% of which are endemic) have been recorded in the Pan-Himalaya region (Chen, 2015). The intercontinental disjunct patterns in *Saussurea* involve nine species: *S. amara* (L.) DC. is broadly distributed in Europe, Asia and North America, and it is the only species found in three continents (Greuter and Raab-Straube, 2006; Keil, 2006; Lipschitz, 1979); *S. alpina* (L.) DC., *S. controversa* DC., *S. parviflora* (Poir.) DC., *S. salsa* (Pall.) Spreng., and *S. turgaiensis* B. Fedtsch. are distributed both in Europe and Asia (Greuter and Raab-Straube, 2006; Lipschitz, 1976, 1979); *S. angustifolia* (L.) DC., *S. nuda* Ledeb. and *S. triangulata* Traut. & C. A. Mey. are distributed both in North America and the northern Asia (Keil, 2006; Lipschitz, 1979). *Saussurea americana* D. C. Eaton and *S. weberi* Hultén are two endemic North American species (Keil, 2006; Lipschitz, 1979). There are three endemic *Saussurea* species in Europe: *S. discolor* (Willd.) DC., *S. porcii* Degen and *S. pygmaea* (Jacq.) Spreng. (Lipschitz, 1979; Greuter and Raab-Straube, 2006;).

The only complete infrageneric classification system of *Saussurea* was proposed by Lipschitz (1979), based on morphological characters. In this system, *Saussurea* was subdivided into six subgenera: subgen. *Jurinocera*, subgen. *Amphilaena*, subgen. *Eriocoryne*, subgen. *Theodorea*, subgen. *Frolovia* and subgen. *Saussurea*.

Susanna and Garcia-Jacas (2007) proposed the *Saussurea* group to include only four genera, i.e., *Saussurea*, *Jurinea* Cass., *Dolomiaea* DC. and *Polytaxis* Bunge. In this treatment, *Hemistepta* Bunge, *Aucklandia* Falc. and *Cavea* W.W.Sm. & Small were included in genus *Saussurea*. Interestingly, *Cavea* was proved to belong to Asteraceae subfamily Gymnarrhenoideae by recent molecular studies (Anderberg et al., 2012; Fu et al., 2016). Recently, the complex of genera *Saussurea*-*Jurinea* was elevated to subtribal rank as subtribe Saussureinae (Herrando-Moraira et al., 2019, in press).

Saussurea (s.l.) as traditionally circumscribed has been considered a highly

polymorphic group (Kita et al., 2004; Raab-Straube, 2003; Wang et al., 2009; Wang et al., 2013). Consequently, based on the results of molecular phylogenies, some species have been recently excluded from *Saussurea* with the aim of circumscribing *Saussurea* (s.s.) as a monophyletic genus and proposing a more natural classification. As a result, two new genera, *Himalaiella* Raab-Straube and *Shangwua* Yu J. Wang et al. were described, and three small generic segregates, *Frolovia* (DC.) Lipsch., *Lipschitzella* Kamelin and *Aucklandia* Falc. were resurrected (Raab-Straube, 2003; Shi and Raab-Straube, 2011; Wang et al., 2013). On the other hand, the genus *Diplazoptilon* was recently merged with *Saussurea* (s.s.) based on evidence from morphological and molecular systematics (Yuan et al., 2015). *Saussurea* sensu stricto is now considered a monophyletic group, with *Polytaxis* and *Hemistepta* as sister groups (Raab-Straube, 2003; Wang et al., 2013; Yuan et al., 2015). In our present study, we adopt the narrow sense concept of genus *Saussurea*. Although the above studies showed that the characters formerly used in *Saussurea* classification are not suitable and shed some light on generic circumscription and intrageneric relationships, the only complete infrageneric classification system of *Saussurea* is still the one established by Lipschitz (1979).

Proposing an infrageneric classification using molecular data is challenging due to extremely low levels of molecular divergence between closely related species and insufficient informative sites (Zhang et al., 2017). This reflects a pattern seen among other members of Asteraceae, which demonstrate substantial morphological variation, but very little molecular differentiation, due to recent and rapid species radiations (Baldwin et al., 1991; Flagel et al., 2008; Wagstaff and Breitwieser, 2004; Zhang et al., 2017). Wang et al. (2009) performed phylogenetic analyses of *Saussurea* based on the nuclear (ITS) and plastid (*trnL-F* and *psbA-trnH*) sequences from 55 species of *Saussurea*, suggested that island-like adaptive radiation occurred during the period of the major uplift events of the Qinghai-Tibetan Plateau, however, without clear infrageneric framework.

Next-generation sequencing (NGS) clearly holds promise for fast and cost-effective generation of multilocus sequence data for phylogeography and phylogenetics (Barrett et al., 2016; Jansen et al., 2005; McCormack et al., 2013; Soltis et al., 2010; Straub et al., 2012), in particular plastome phylogenies. The idea of using whole chloroplast genomes to identify plant species was first proposed by Kang and Cronk (2008) and has been highlighted by a few recent review articles not only for identifying plant species but also for distinguishing related groups (Li et al., 2015; Liang et al., 2019; Parks et al., 2009; Zhang et al., 2017). In this study we use complete chloroplast genome sequence data on a comprehensive sample of species and geographical areas of *Saussurea*, to provide insights into its infrageneric classification, geographic history, and also to explore the potential informative characters for this genus.

2. Materials and methods

1 2.1. Taxon sampling

2 A total of 126 species of *Saussurea* (s.s.) were collected from China, Russian Siberia,
3 Japan, Korea, Europe and North America to represent the global distribution area of the
4 genus. Our sampling covered all subgenera and most sections of *Saussurea* according
5 to either the infrageneric system of Lipschitz (1979), Shi and Raab-Straube (2011), and
6 Chen (2015). Sixteen additional samples of species of *Dolomiaea*, *Himalaiella*,
7 *Bolocephalus*, *Aucklandia*, *Cirsium*, *Arctium*, *Carthamus*, *Centaurea*, *Silybum*,
8 *Atractylodes* were also chosen based on previous studies (Kita et al., 2004; Raab-
9 Straube, 2003; Y. J. Wang et al., 2009). Plant materials were mostly collected in the
10 field, three were obtained from herbarium (A) collections (*S. americana*, *S. nuda* and
11 *S. alpina*). Collection details of the specimens are shown in Appendix Table 1.

12 2.2. DNA extraction and sequencing

13 Total genomic DNA was extracted from silica gel-dried tissue or herbarium tissue of
14 one plant per population using CTAB Plant Genomic DNA Extraction kit (Biomed
15 BeiJing). The quantity of each extraction was checked with Qubit™ 2.0 Fluorometer
16 (Thermo Scientific, Waltham, MA, USA). Around 400 ng of DNA was sonicated using
17 a Covaris S2 (Covaris, Woburn, MA) to produce fragments ~150-350 bp in length for
18 making sequence libraries for paired-end reads. To ensure that genomic DNA was
19 sheared at approximately the selected fragment size, all samples were checked and
20 evaluated on a 1.2% (w/v) agarose gel. Library construction was performed with
21 NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® (New England Biolabs,
22 MA, USA) following the manufacturer's suggested procedures. Sequencing was done
23 on one lane with illumina Hiseq 2500 sequencing platform (Illumina, San Diego,
24 California, USA) in Tianjin Novogene Bioinformatics Technology Company LTD
25 using 100 bp paired-end reads. The expected sequencing quantity of each sample is 4
26 Gigabyte, and the number of Raw FASTQ data reads is shown in Appendix Table 1.

27 2.3. Raw data processing

28 A first quality control of raw sequence reads demultiplexed by sequencing scores was
29 conducted in FastQC v.0.10.1
30 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Removing of Illumina
31 adapters and cleaning of Raw FASTQ data was then carried out using ApterRemoval
32 v2 (Lindgreen, 2012) and Trimmomatic (Bolger et al., 2014). Trimming step was
33 conducted with a slidingwindow set to 1:5, cutting a read when the average of five
34 examined positions falls below 20 of the quality Phred+33 score. Cleaned reads finally
35 retained were those with a minimum length of 50 bp and with both corresponding
36 forward and reverse pair.

37 2.4. Sequence assembly and annotation

38 Geneious 11.0.2 (Kearse et al., 2012) was used to assemble all sequence using

Saussurea chabyoungsanica Im (KX622799) as reference. Firstly, cleaned paired-end reads mapped to reference with medium-low sensitivity in three iterations, the mapped reads were assembled de novo with Velvet 1.2.10 (Zerbino and Birney, 2008) setting kmer length to 89. Secondly, the contigs obtained were then combined and imported as reference to extend, Geneious were used to assemble the second round contigs. Repeat mapping using the contigs from the previous round as references until all contigs were assembled to one, then marked and integrated large single region, repeat region and small single region to complete plastomes. The consensus sequence was annotated using *S. chabyoungsanica* (KX622799) as references in Geneious 11.0.2 and then corrected manually. The tRNA regions were confirmed using tRNA scan-SE version 1.21 (Schattner et al., 2005) with default settings. All region including coding and uncoding region were marked with the .bed file generated by Geneious. The last repeating region (IRB) was removed from the sequences and the rest regions were extracted according to the mark individually.

2.5. Sequence alignment and phylogenetic analyses

Sequences were aligned using Mafft-win v7.221 (Katoh and Standley, 2013; Yamada et al., 2016), coding regions using codon alignment mode and uncoding region by default, then adjusted manually in BioEdit 7.1.8 (Hall, 1999), the regions of length less than 20bp were deleted. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis was done with Gblocks 0.91b (Castresana, 2000) and trimAl v.14 (Capella-Gutiérrez et al., 2009) in default setting. DAMBE 6 (Xia, 2018) was used to measure substitution saturation by default. Gene alignments were concatenated with PhyloSuite (Zhang et al., 2018) for subsequent phylogenetic. Model selection for each alignments were processed in PartitionFinder 2 (Calcott et al., 2016) by default. Finally, summary statistics of concatenated matrices were computed with AMAS (Borowiec, 2016).

Cirsium arvense, *Cirsium eriophorum*, *Cirsium japonicum* var. *maackii*, *Cirsium japonicum* var. *spinosissimum*, *Cirsium vulgare* and *Silybum marianum* were chosen as the outgroups and genus *Atractylodes* DC. was excluded in phylogeny analyze. Bayesian phylogenetic inference (BI) analysis was conducted with MrBayes 3.2.6 (Ronquist and Huelsenbeck, 2003) in Cipres. The partition and model selection resulting from PartitionFinder 2 were used. Two independent analyses consisting of four Markov Chain Monte Carlo (MCMC) chains were run for 10 million generations, sampling one tree every 1000 generations, runs were completed when the average standard deviation of split frequencies reached 0.01, the stationarity of the runs was assessed using Tracer v1.7 (Rambaut et al., 2018), after removing the burn-in period samples (the first 25% of sampled trees), a 50% majority rule consensus tree was constructed. Maximum Likelihood (ML) analysis was performed with RAxML-HPC v.8 (Stamatakis, 2014) on XSEDE also in Cipres portal (Miller et al., 2010). The partition result of PartitionFinder 2 were used. GTR+G model was chosen for bootstrapping and we also ran a simultaneous rapid bootstrapping and best ML tree search (Stamatakis et al., 2008), with 10 randomized maximum parsimony starting trees

and a bootstrap resampling of 1000 replicates to assess branch support values. The Bayesian posterior probability (PP) from BI analysis and bootstrap values (BS) from ML of each branch was obtained. Nodes with $PP \geq 0.95$ (Ronquist and Huelsenbeck, 2003) and $BS \geq 75$ (Hillis and Bull, 1993) were considered well supported.

2.6. Molecular age estimation

Dating analysis were conducted in treePL (Smith and O'Meara, 2012) using the RAxML tree with genus *Atractylodes* included as outgroup. TreePL estimates evolutionary rates and divergence dates on a tree given a set of fossil constraints and a smoothing factor determining the amount of among branch rate heterogeneity (Pyron, 2014; Smith and O'Meara, 2012). We chose TreePL rather than Beast or MrBayes because of it is computationally tractable for our data of such size while the result of age estimation are almost the same in former studies (Klimov et al., 2017; Stubbs et al., 2018). Analyses were not only based on the fossil calibration points but also the second calibration points, given the computational difficulties of dating trees of this size using minimum and maximum ages (Bremer et al., 2007). Minimum age was based on former fossil studies (Harrison et al., 1992; Mai and Hans, 1995; Mai and Dieter, 2001; Popescu, 2002) and maximum based on dating result of *Cardueae* (Laia Barres et al., 2013) based on 4 chloroplast regions. One second calibration (A) and three fossil calibration points (B, C, D) and were used: A. The crown age of *Carduae* was set to 39.2 Mya (95%HPD: 35.03-43.49) based on the estimation by Barres et al. (2013). B. Stem lineage of *Carduus* - *Cirsium* was set to a minimum age of 14 Ma based on Miocene achenes identified as *Cirsium* (Mai and Hans, 1995) and maximum age to 18.26 Mya. C. The minimum stem age for *Centaurea* was set to 6 Ma, based on several records of pollen and achenes for this genus dating from the early Pliocene onward (Mai and Hans, 1995; Popescu, 2002) and maximum age to 19.69 Mya. D. Split of lineages *Arctium-Cousinia* group and *Jurinea-Saussurea* (Barres et al., 2013; Wang et al., 2009) group was set to a minimum age to 8 Ma, based on mid-late Miocene achenes described by Mai and Dieter (2001) and assigned to *Arctium* (López-Vinyallonga et al., 2009) and maximum age to 29.24 Mya.

A 'priming' analysis was first conducted to determine the best optimization parameters. Based on the results of 100 repeats, the values of gradient-based optimizer were set to 2, autodiff based optimizer to 2, and autodiff cross-validation based optimizer to 1. The random subsample and replicate cross-validation (RSRCV) analyses were conducted from 0.001 to 10000 (in 10-fold increments) to determine the best smoothing value (which was found to be 0.001 after 1000 repeats), the number of cross validation and penalized likelihood total optimization iterations was set to 4, the number of cross validation and penalized likelihood simulated annealing iterations was set to 10000.

2.7. Ancestral geographical area reconstruction

Ancestral range reconstruction was conducted to estimate possible historical patterns of geographical distribution across *Saussurea*. We defined five geographical areas (Fig.

1) based on the Vegetation Map of the People's Republic of China (1:1000000) (Zhang, 2007), the revised Thorthwaite-type climate classification (Feddema, 2005), Floristic regions of the world (Thorne, 1987) and the distribution of *Saussurea* (<http://www.cvh.ac.cn/>; http://www.efloras.org/flora_page.aspx?flora_id=1; <http://ww2.bgbm.org/EuroPlusMed/query.asp>; <https://plant.depo.msu.ru/?d=p>): A, Hengduan Mountains; B, Qinghai-Tibetan Plateau; C, Central Asia, Northwest China and Mongolia; D, Southern to Northeastern China, Japan & Korean Peninsula and Russian Far East; E, Europe, North and mid-west Asia and North America. A is the diversity center of *Saussurea*, B is the alpine and plateau area, C is the warm temperate desert region, D is the low altitude forest region in east Asia and E is other distribution areas in north temperate region. BioGeoEARS package (Matzke, 2013) in R were used to study the ancestral geographical area of *Saussurea* based on the RaXML tree dated by TreePL with *Saussurea* species only According to the instructions (<http://phylo.wikidot.com/biogeobears#toc3>). We tested all three models: DEC, BAYESAREALIKE, and DIVALIKE. Additionally, a parameter (+J) for founder-event speciation was included, which is likely important in montane systems. We used likelihood-ratio tests and AIC (Akaike Information Criterion) values to compare the fit of these models to the data and DEC+J was selected.

3. Results

3.1. Chloroplast genomes of *Saussurea*

Chloroplast genomes of 126 *Saussurea* and 7 outgroup species were got from our work and others were download from GenBank. GenBank accession numbers and length of each chloroplast genomes are listed in Appendix Table 1. The length of *Saussurea* chloroplast genomes ranged from 152024 bp (*S. sobarocephala*) to 152892 bp (*S. uliginosa*). Chloroplast genome structure is very conservative, it not only reflects in the size of whole chloroplast genome structure, but the order and size of each gene and intergenic region.

3.2. Phylogenetic analyses

BI and ML analyses resulted in identical tree (Fig. 2) based on chloroplast genomes data matrix, 140 regions, total length 126538bp. All *Saussurea* samples form a monophyly with genus *Hemistepta* as sister group, then divided in three clades. However, our three clades tree could not completely support the classification system of *Saussurea* provided by Lipschitz (1979), although after moving in or out some groups (Kita et al., 2004; Raab-Straube, 2003; Shi and Raab-Straube, 2011; Wang et al., 2013; Yuan et al., 2015). Clade 1 (PP = 1, BS = 96) includes three species sampled of subgen. *Amphilaena* (Stsch.) Lipsch., one species of subgen. *Eriocoryne* (Wall. ex DC.) Hook. f., eight species of subgen. *Theodorea* Cass. and forty-five species of subgen. *Saussurea*; Clade 2 (PP = 1, BS = 100) includes four species sampled of subgen. *Amphilaena*, eight species of subgen. *Eriocoryne*, and 13 species of subgen. *Saussurea*; Clade 3 (PP = 1, BS = 91) includes eight species sampled of subgen. *Amphilaena*, four

1 species of subgen. *Eriocoryne*, and 28 species of subgen. *Saussurea*.

2 Previous molecular studies not only do not support the monophyly of Lipschitz's
3 concept of *Saussurea*, but also do not support the monophyly of nearly all the subgenera.
4 Some morphological traits such as translucent bracts which was used to define
5 subgenus *Amphilaena* is proved to have evolved in parallel evolution (Raab-Straube,
6 2003; Wang et al., 2009; Yuan et al., 2015). A well-supported infrageneric framework
7 of *Saussurea* based on molecular data is still lacking.

8 In our present study, all samples from subgenus *Theodorea*, subgenus *Saussurea*
9 section *Laguranthera* and section *Rosulascentes* are in clade 1; 13 of 17 sampled
10 species in subgen. *Eriocoryne* are in clade 2.

11 To study which and how many sequences could provide enough information, we
12 recorded and analyzed variable sites, proportion of variable sites, parsim-informative
13 sites and proportion of parsim-informative sites of each region (Appendix Table 2) and
14 the distribution are shown on the Appendix Fig.1. The number of variable sites in repeat
15 region are less than large single copy (LSC) and small single copy (SSC) shows repeat
16 region is more conservative (Palmer and Thompson, 1982; Shaw et al., 2007). The
17 number of variable sites in uncoding region (more than 20 bp) is 7752 higher than
18 coding regions, 3216. Present study also made phylogeny analyses using coding region
19 only (Appendix Fig 2), however, the topology structure are instability and the supports
20 value are lower than using all regions. Including uncoding regions into matrix is a good
21 method to get a better result when using whole chloroplast genomes(Dong et al., 2012).
22 30 most informative regions (Proportion-parsimony informative-sites) are needed at
23 least to get a well-supported ML tree (Appendix Fig. 3). These 30 regions have 1994
24 parsim-informative sites, about 38% of the whole (5194 sites), all the gene sequences
25 of former authors (Wang et al., 2009) used have 94 parsim-informative sites at most,
26 less than 2% of the whole. It could explain why former studies could not get a clear
27 result using few gene sequences.

28 3.3. Divergence time estimation and biogeographical reconstruction

29 Divergence time estimates for *Saussurea* are shown in Fig.2 and ancestral area
30 reconstructions in Fig.3. *Saussurea* separated with *Hemistepa* about 18.54 Mya and
31 started its diversification about 17.05 Mya at Hengduan Mountains (A area) and then
32 diffused around and dispersed northward. Clade 1 includes the main one northward
33 clade including about 90% species distribution in north area (C, D and E area), however,
34 the origin area is also in Hengduan Mountains. The most recent common ancestor of
35 northward species occupied in warm temperate desert region (C area) in about 16.31
36 Mya. Clade 2 and 3 mainly distribute in A and B area, but also include some species
37 dispersal northward (*S. globosa*, *S. katochaete* and *S. graminea* in clade 2; *S.*
38 *involucrata*, *S. orgaadayi*, *S. tangutica*, *S. sutchuenensis* and *S. subulata* in clade 3).

39

4. Discussion

4.1. Phylogenetic relationships within *Saussurea*

Up to now, a lack of a well-supported infrageneric framework based on molecular data is a big problem on the classification of *Saussurea*. Analyses based on a few chloroplastic markers could not provide enough resolution due to low levels of molecular divergence caused by the rapid radiation of the genus (Wang et al., 2009; Yuan et al., 2015). Whole chloroplast genome sequencing is a better method to study a genus like *Saussurea* rather than several short markers, due to its low consumption and high support values.

Although chloroplast genome data give a pattern of the infrageneric framework of *Saussurea* and some evidence also support it, more convincing evidence is still required. Nuclear genes, cytological and morphology evidence especially are needed to further understand, due to the criticism to the use of plastome for unraveling closely related species relationships: hybridization, incomplete lineage sorting, chloroplast capture and maternal inheritance may bias phylogenetic relationships (Hörandl and Appelhans, 2015; Samuel et al., 2019, Herrando-Moraira et al., 2019). Recently, Herrando-Moraira et al. (2018) give us a different 3-clades pattern using nuclear markers in NGS target enrichment method, which offers a new sight to study the phylogenetic relationships within *Saussurea*, although some important areas, such as QTP and Hengduan Mountains, were not represented in this study.

Some groups of *Saussurea* show close relationships in our phylogeny, which constitutes a good suggestion for further studying the classification of *Saussurea*. The relationships of subgenus *Theodorea*, subgenus *Saussurea* section *Laguranthera* and section *Rosulascentes* is a good example: not only chloroplast genomes demonstrate the closely relationships of these three groups, but morphology and geography support this: middle bracts margin shared obvious colored appendages in these three group, while other clades share other form of margin appendages, or lack appendages; and almost all of these three groups are distributed in areas beyond the alpine plateau (C, D and E area, Fig. 1). Actually, a close relationship of these groups was pointed out in previous studies: eight species from these groups formed a clade in the parsimony tree based on *trnL-F* sequences by Wang and Liu (2004); and nine species formed a clade in the ML tree based on *trnL-F*, *psbA-trnH*, and ITS sequences in Wang et al. (2009). Raab-Straube (2017) also mentioned the close relationship of subgenus *Theodorea*, subgenus *Saussurea* section *Laguranthera*. Another example is the species of subgenus *Eriocoryne*: most of them are placed in clade 2 and almost all of them are distributed in the high altitudes of the Hengduan Mountains and QTP area.

4.2. Historical biogeography of *Saussurea*

According to our results, *Saussurea* diverged at about 18.54 Mya, however, some studies also work on the divergence time had different result. Wang et al. (2009)

1 included 55 *Saussurea* species in their dating analysis, a much younger estimate than
2 present study got, 11.7–14.4 Mya based on the nuclear internal transcribed spacer (ITS)
3 dataset and 8.1–9.7 Mya derived from the combined nuclear internal transcribed spacer
4 (ITS) and plastid *trnL-F* and *psbA-trnH* sequences. The calibrated the crown age of the
5 tribe Cardueae based on Kim et al. (2005) and the pollen fossil of *Cirsium* (≥ 5 Mya,
6 Menke, 1976) used as a calibration point, instead of the earliest fossil record known for
7 this genus, the mid Miocene achenes described in Mai and Hans (1995), could explain
8 the younger result. There are two studies focus on the origin of tribe Cardueae
9 (Compositae) and also estimated the divergence time of *Saussurea*. One is the study of
10 Barres et al. (2013), including two *Saussurea* species based on four chloroplast regions
11 (*trnL-trnF*, *matK*, *ndhF*, *rbcL*), the estimated stem age of *Saussurea* is older than our
12 result, 19.96 Mya. Another one is 12.48 Mya, raising by Herrando-Moraira et al. (2019),
13 younger than our estimated, based on nuclear Hyb-Seq data, three *Saussurea* species
14 included. They improved the fossil calibration points and three of them in tribe
15 Cardueae were employed by this study. The result of these two studies different from
16 ours is mainly due to sampling (Linder et al., 2005, Herrando-Moraira et al., 2019), so
17 few *Saussurea* species included and the most closely related outgroup is *Cousinia* or
18 *Jurinea* rather than *Hemistepta* that we used. Different dataset is another interesting
19 thing, our matrix is based on chloroplast and the result is more close to Barres et al.
20 (2013) which is also based on four chloroplast regions, whereas the data Herrando-
21 Moraira et al. (2019) used was nuclear sequences and their result is more close to Wang
22 et al. (2009) which used ITS. Different data, sampling and method could have impact
23 on the dating result, just as Herrando-Moraira et al. (2019) mentioned, more explicit
24 methods to address these processes should be performed, in addition to using a wider
25 species sampling.

26 The uplift of QTP is an important event in geological history since the Cenozoic playing
27 an important role in the change of geography, climate, vegetation on the QTP and
28 adjacent areas (Wen et al., 2014). The origin and evolution of many North Temperate
29 plant groups were influenced by the uplift of QTP: *Nannoglottis* (Liu et al., 2002),
30 *Caragana* (Zhang and Fritsch, 2010), *Rheum* (Sun et al., 2012) and *Rhodiola* (Zhang et
31 al., 2014). Regarding the origin of *Saussurea*, Wang et al. (2009) studied the
32 distribution pattern of this genus and hypothesized that the group probably originated
33 in the Central Asia-Himalaya region, and the ancestral, primitive *Saussurea* might have
34 originated during the Tertiary period. Raab-Straub (2017) raised that the *Saussurea*
35 clade probably began to develop in Central Asia, in the Tian Shan and the Altai regions,
36 based on analyzing of relation groups and chromosome number. Our dating analysis,
37 however, suggest that *Saussurea* was most likely originated in the Hengduan Mountains
38 during the early-middle Miocene (18.54 Mya).

39 The distribution of *Saussurea* species is closely related to the uplift of QTP, and the
40 main reason may be the barrier of west and middle QTP. Some areas of west and middle
41 QTP had reached about 5000 m and formed barrier before the *Sausurea* appeared, for
42 instance: Kangrinboqe area of Lhasa block reached 4700–6700 m at about 24 Mya
43 (DeCelles et al., 2011), Heihuling area reached 4600–5800 m during 58–28 Mya (Xu

et al., 2013) and Linzhou basin reached 4050–4950 m during 50–60 Mya (Ding et al., 2014). Geographic isolation led to the arrival of *Saussurea* in Hengduan Mountains area where the altitude was lower and the environment was suitable at about 18.54 Mya (Wang et al., 2006; Rowley and Garziona, 2007; Polissar et al., 2009; Xu et al., 2010; Gébelin et al., 2013; Favre et al., 2015; Sun et al., 2015;).

There are two possible main reasons for the diffusion of the *Saussurea*: continuous changing of the QTP, on the one hand, and global cooling from Mid-Miocene, on the other. The QTP orogeny during the Pliocene and Pleistocene created a vast array of new habitats across wide elevational ranges (up to 5000 m in the Hengduan Mountains, which are perhaps the largest ‘evolutionary front’ of the world’s temperate zone) and stimulated allopatric and habitat differentiation, thus ultimately giving rise to adaptive radiations (Liu and Tian, 2007; López-Pujol et al., 2011) leading to the rapid diversification of *Saussurea* in the region. In only about 2 Mya *Saussurea* was divided into three clades in Hengduan Mountains; then, most of *Saussurea* (almost all clade 2 and clade 3, and part of clade 1) diffused around QTP and Hengduan Mountains. As a result, QTP and Hengduan Mountains have the greatest diversity of the genus (Chen, 2015; Wang et al., 2009). The global cooling since mid-Miocene about 16 Mya (Zachos et al., 2001) is another important factor in spreading of *Saussurea*. It resulted in lower and northward dispersal, and the northward clades (stem age ca. 16.97 Mya) in clade 1 are a persuasive evidence of this extension.

Debate continues over whether global cooling or uplifting of the Tibetan Plateau provided the first-order driver for the aridification of Central Asia throughout the Mid-Late Miocene (Miao et al., 2012). However, it is obvious that the uplift of the Tibetan Plateau led to the arising of *Saussurea* and the continuous changing QTP and global cooling given the opportunity to spread, global cooling since mid-Miocene accelerating especially northward spreading.

The clear patterns of how the changing of the QTP and global cooling results in *Saussurea* evolution is not simple. Almost every area was composed of species from different clades in different time. There are 13 species from clade 1, 19 species from clade 2 and 30 species from clade 3 in A area. It suggests that there are at least three independent colonizations of the Hengduan Mountains. The same condition also occurred in other areas (B area, 6 species from clade 1, 8 species from clade 2 and 17 species from clade 3; C area: 35 species from clade 1, 7 species from clade 2 and 12 species from clade 3; D area: 21 species from clade 1, 2 species from clade 2, 4 species from clade 3). This is the same at the origin of European and North American species: neither samples from Europe nor North America could form a single clade. They are mixed with the north Asian species and formed parallel clades: for instance, *S. alpina* and *S. controversa* from Europe are parallel with most species from north Asia, *S. nuda* from North America and *S. pseudoalpina* from northwestern China form a clade and parallel with other species from North America and North China.

Wen et al. (2014) summarized six patterns of evolutionary diversifications of plants on

the QTP, and three of them at least occurred in *Saussurea*. (1). The QTP as a biogeographic source area in Eurasia (Jia et al., 2012; Wang et al., 2004), also called out-of-QTP hypothesis (Fan et al., 2015; Wang and Liu, 2004). As our result suggested, all the three clades originated in the Hengduan Mountains and then expanded to the Northern China, Europe and North America. (2). The QTP as a major biogeographic barrier for plant diversification in Eurasia (Xie et al., 2014; Yi et al., 2008) is the second pattern. There are few *Saussurea* species in west and south of the Himalayas whereas most species are found in another side. (3). The origin area of northward clades in clade 1 were Central-North Asia, and some species spread back to Hengduan Mountains area, e.g. *S. peduncularis*, *S. malitiosa* and *S. tsoongii*. This pattern is similar to the radiation of genus *Maianthemum* on the QTP following the migration from the Sino-Japanese floristic region (Meng et al., 2008). More samples and more evidences are needed, because we only have about 25% samples of all *Saussurea* after all. A more comprehensive study will not only give clear patterns on how *Saussurea* developed, but also on the evolutionary diversification of plants on the QTP.

Saussurea brachycephala from Japan is closely related to the species from the Korean Peninsula and Northeastern China, which reflect the North-China origin of Japanese species. Most taxa of *Saussurea* from North America are closely related to arctic taxa from Siberia and Far East of Russia. This relationship suggests that *Saussurea* might spread from Siberia and Far East of Russia into North America through the Bering land bridge. This kind of migration pattern has also been found in another Asteraceae genus from tribe Cardueae, *Plectocephalus* (Susanna et al., 2011).

5. Conclusions

A well-supported 3-clades phylogenetic framework of *Saussurea* was established based on whole chloroplast genomes. However, our framework is not consistent with former infrageneric classification on the whole genus even after moving in or out some groups based on recent studies. Some groups show certain relationships: subgenus *Eriocoryne* was almost identified, and subgenus *Theodorea*, subgenus *Saussurea* section *Laguranthera* and section *Rosulascentes* are closely related.

Saussurea originated in Hengduan Mountains about 18.54 Mya and diversified quickly in this area, and dispersed northwards since then. The barrier effect of QTP maybe the main reason for current distribution of *Saussurea*. Continuous changing of QTP and global cooling since mid-Miocene should be responsible for diversification and spreading, and especially the later, for the northwards spread of *Saussurea*.

Acknowledgements

We are grateful to curators of the herbaria of A, E, K, KUN and PE for allowing us to study their collections or for sending specimens on loan. We thank Ms. M. N. Shurupova, Dr. Z. H. Wang, Mr. L. Yuan and Ms. T. J. Tong for collecting materials, Dr. D. E. Boufford for sending materials, Mercè Galbany-Casals and anonymous

reviewer for his invaluable comments on the manuscript. This research was supported by the National Natural Science Foundation of China (grant nos.: 31670190, 31370226).

References

- Anderberg, A.A., Ohlson, J. L., 2012. The genus *Cavea*, an addition to the tribe Gymnarrheneae (Asteraceae–Gymnarrhenoideae). *Compositae Newsletter* 50: 46–55.
- Baldwin, B.G., Kyhos, D.W., Dvorak, J., Carr, G.D., 1991. Chloroplast DNA evidence for a North American origin of the Hawaiian silversword alliance (Asteraceae). *Proc. Natl. Acad. Sci. U. S. A.* 88, 1840–3. <https://doi.org/10.1073/PNAS.88.5.1840>
- 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>
- Barrett, C.F., Bacon, C.D., Antonelli, A., Cano, Á., Hofmann, T., 2016. An introduction to plant phylogenomics with a focus on palms. *Bot. J. Linn. Soc.* 182, 234–255. <https://doi.org/10.1111/boj.12399>
- 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., Jacquet, D., Lundqvist, S., Britton, T., Anderson, C.L., 2007. Estimating divergence times in large phylogenetic trees. *Syst. Biol.* 56, 741–752. <https://doi.org/10.1080/10635150701613783>
- Calcott, B., Lanfear, R., Senfeld, T., Frandsen, P.B., Wright, A.M., 2016. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, msw260. <https://doi.org/10.1093/molbev/msw260>
- 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>
- 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>
- Chen, Y.S., 2015. Flora of Pan-Himalaya, vol. 48(2), Asteraceae II: *Saussurea*. Science Press, Beijing; Cambridge university Press, Cambridge.
- DeCelles, P.G., Kapp, P., Quade, J., Gehrels, G.E., 2011. Oligocene-Miocene Kailas basin, southwestern Tibet: Record of postcollisional upper-plate extension in the

1 Indus-Yarlung suture zone. Geol. Soc. Am. Bull. 123, 1337–1362.
2 <https://doi.org/10.1130/B30258.1>

3 Ding, L., Xu, Q., Yue, Y., Wang, H., Cai, F., Li, S., 2014. The Andean-type Gangdese
4 Mountains: Paleoelevation record from the Paleocene–Eocene Linzhou Basin.
5 Earth Planet. Sci. Lett. 392, 250–264. <https://doi.org/10.1016/j.epsl.2014.01.045>

6 Dong, W., Liu, J., Yu, J., Wang, L., Zhou, S., 2012. Highly variable chloroplast markers
7 for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding.
8 PLoS One 7, 1–9. <https://doi.org/10.1371/journal.pone.0035071>

9 Fan, J.Y., Chen, H. B., Zhu, L., Chen, H.L., Zhao, Z.-Z., Yi, T., 2015. *Saussurea medusa*, source
10 of the medicinal herb snow lotus: a review of its botany, phytochemistry,
11 pharmacology and toxicology. Phytochem. Rev. 14, 353–366.
12 <https://doi.org/10.1007/s11101-015-9408-2>

13 Favre, A., Michalak, I., Chen, C.-H., Wang, J.-C., Pringle, J.S., Matuszak, S., Sun, H., Yuan, Y.-M., Struwe, L.,
14 Muellner-Riehl, A.N., 2016. Out-of-Tibet: the spatio-temporal evolution of
15 *Gentiana* (Gentianaceae). J. Biogeogr. 43, 1967–1978.
16 <https://doi.org/10.1111/jbi.12840>

17 Favre, A., Päckert, M., Pauls, S.U., Jähnig, S.C., Uhl, D., Michalak, I., Muellner-Riehl,
18 A.N., 2015. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution
19 of Tibetan biotas. Biol. Rev. 90, 236–253. <https://doi.org/10.1111/brv.12107>

20 Feddema, J.J., 2005. A revised thornthwaite-type global climate classification. Phys.
21 Geogr. 26, 442–466. <https://doi.org/10.2747/0272-3646.26.6.442>

22 Flagel, L.E., Rapp, R.A., Grover, C.E., Widrlechner, M.P., Hawkins, J., Grafenberg,
23 J.L., Alvarez, I., Chung, G.Y., Wendel, J.F., 2008. Phylogenetic, morphological,
24 and chemotaxonomic incongruence in the North American endemic genus
25 *Echinacea*. Am. J. Bot. 95, 756–765. <https://doi.org/10.3732/ajb.0800049>

26 Gébelin, A., Mulch, A., Teyssier, C., Jessup, M.J., Law, R.D., Brunel, M., 2013. The
27 Miocene elevation of Mount Everest. Geology 41, 799–802.
28 <https://doi.org/10.1130/G34331.1>

29 Greuter, W., Raab-Straube, E. Von, 2006. Euro+Med Notulae, 2. Willdenowia 36, 707–
30 717. <https://doi.org/10.3372/wi.36.36206>

31 Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment program for
32 Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

33 Harrison, T.M., Copeland, P., Kidd, W.S.F., Yin, A., 1992. Raising Tibet. Science
34 255, 1663–1670. <https://doi.org/10.1126/science.255.5052.1663>

35 Herrando-Moraira, S., Calleja, J.A., Carnicero, P., Fujikawa, K., Galbany-Casals, M.,
36 Garcia-Jacas, N., Im, H.T., Kim, S.C., Liu, J.Q., López-Alvarado, J., López-Pujol,
37 J., Mandel, J.R., Massó, S., Mehregan, I., Montes-Moreno, N., Pyak, E., Roquet,
38 C., Sáez, L., Sennikov, A., Susanna, A., Vilatersana, R., 2018. Exploring data
39 processing strategies in NGS target enrichment to disentangle radiations in the
40 tribe Cardueae (Compositae). Mol. Phylogenet. Evol. 128, 69–87.
41 <https://doi.org/10.1016/j.ympev.2018.07.012>

- 1 Herrando-Moraira, S., Calleja, J.A., M., Garcia-Jacas, N., Liu, J.Q., López-Alvarado,
2 J., López-Pujol, J., Mandel, J.R., Massó, S., Montes-Moreno, N., Roquet, C., Sáez,
3 L., Sennikov, A., Susanna, A., Vilatersana, R., 2019. Nuclear and plastid DNA
4 phylogeny of the tribe Cardueae (Compositae) with Hyb-Seq data: A new subtribal
5 classification and a temporal diversification framework. *Mol. Phylogenet. Evol.*
6 (in press)
- 7 Hörandl, E., Appelhans, M.S., 2015. Next Generation Sequencing in Plant Systematics.
8 Koeltz Botanical Books, Oberreifenberg.
- 9 Jansen, R.K., Raubeson, L.A., Boore, J.L., DePamphilis, C.W., Chumley, T.W.,
10 Haberle, R.C., Wyman, S.K., Alverson, A.J., Peery, R., Herman, S.J., Fourcade,
11 H.M., Kuehl, J. V., McNeal, J.R., Leebens-Mack, J., Cui, L., 2005. Methods for
12 obtaining and analyzing whole chloroplast genome sequences. pp. 348–384. in:
13 Zimmer, E.A., Roalson, E.H. (Eds.), *Methods in Enzymology*. Academic Press ,
14 Cambridge, Massachusetts. [https://doi.org/10.1016/S0076-6879\(05\)95020-9](https://doi.org/10.1016/S0076-6879(05)95020-9)
- 15 Jia, D.-R., Abbott, R.J., Liu, T.-L., Mao, K.-S., Bartish, I. V., Liu, J.-Q., 2012. Out of
16 the Qinghai-Tibet Plateau: evidence for the origin and dispersal of Eurasian
17 temperate plants from a phylogeographic study of *Hippophaë rhamnoides*
18 (Elaeagnaceae). *New Phytol.* 194, 1123–1133. [https://doi.org/10.1111/j.1469-](https://doi.org/10.1111/j.1469-8137.2012.04115.x)
19 [8137.2012.04115.x](https://doi.org/10.1111/j.1469-8137.2012.04115.x)
- 20 Kane, N.C., Cronk, Q., 2008. Botany without borders: barcoding in focus. *Mol. Ecol.*
21 17, 5175–5176. <https://doi.org/10.1111/j.1365-294X.2008.03972.x>
- 22 Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software
23 version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–
24 780. <https://doi.org/10.1093/molbev/mst010>
- 25 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton,
26 S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P.,
27 Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop
28 software platform for the organization and analysis of sequence data.
29 *Bioinformatics* 28, 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- 30 Keil, D.J., 2006. *Saussurea*. Pp: 165–168. In: Barkley, T.M., Brouillet, L., Strother, J.
31 L. (eds.) *Flora of North America*. Vol. 19. Oxford University Press, New York.
- 32 Kita, Y., Fujikawa, K., Ito, M., Ohba, H., Kato, M., 2004. Molecular phylogenetic
33 analyses and systematics of the genus *Saussurea* and related genera (Asteraceae,
34 Cardueae). *Taxon* 53, 679. <https://doi.org/10.2307/4135443>
- 35 Klimov, P.B., Mironov, S. V., OConnor, B.M., 2017. Detecting ancient codispersals
36 and host shifts by double dating of host and parasite phylogenies: Application in
37 proctophyllodid feather mites associated with passerine birds. *Evolution* 71, 2381–
38 2397. <https://doi.org/10.1111/evo.13309>
- 39 Li, X. W., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y.T., Chen, S. L., 2015. Plant
40 DNA barcoding: from gene to genome. *Biol. Rev.* 90, 157–166.

1 <https://doi.org/10.1111/brv.12104>

2 Li, X.W., Li, J., 1993. A preliminary floristic study on the seed plants from the region
3 of Hengduan Mountain. *Acta Bot. Yunnan.* 15, 217–231.

4 Liang, J. Y., Ma, Q., Yang, Z. P., 2019. The first complete chloroplast genomes of two
5 Alismataceae species, and the phylogenetic relationship under order Alismatales.
6 Mitochondrial DNA Part B 4, 122–123.
7 <https://doi.org/10.1080/23802359.2018.1536486>

8 Linder, H.P., Hardy, C.R., Rutschmann, F., 2005. Taxon sampling effects in molecular
9 clock dating: An example from the African Restionaceae. *Mol. Phylogenet. Evol.*
10 35, 569–582. <https://doi.org/10.1016/j.ympev.2004.12.006>

11 Lindgreen, S., 2012. AdapterRemoval: Easy cleaning of next-generation sequencing
12 reads. *BMC Res. Notes* 5. <https://doi.org/10.1186/1756-0500-5-337>

13 Lipschitz, S., 1976. *Saussurea* DC. Pp: 216–217. in: Tutin, T. G., Heywood, V.H.,
14 Burges, N.A., Moore, D. M., Valantine, D. H., Walters, S.M., Webb, D.A. (eds.)
15 *Flora Europaea*, volume 4. Cambridge University Press, Cambridge.

16 Lipschitz, S., 1979. Genus *Saussurea* DC. (Asteraceae). Nauka, Leningrad. (in Russian).

17 Liu, J.-Q., Gao, T.-G., Chen, Z.-D., Lu, A.-M., 2002. Molecular phylogeny and
18 biogeography of the Qinghai-Tibet Plateau endemic *Nannoglottis* (Asteraceae).
19 *Mol. Phylogenet. Evol.* 23, 307–325. [https://doi.org/10.1016/S1055-](https://doi.org/10.1016/S1055-7903(02)00039-8)
20 7903(02)00039-8

21 López-Vinyallonga, S., Mehregan, I., Garcia-Jacas, N., Tscherneva, O., Susanna, A.,
22 Kadereit, J.W., 2009. Phylogeny and Evolution of the *Arctium-Cousinia* Complex
23 (Compositae, Cardueae-Carduinae). *Taxon* 58: 153–171.
24 <https://doi.org/10.2307/27756831>

25 Mai, D.H., Hans, D., 1995. Tertiäre Vegetationsgeschichte Europas : Methoden und
26 Ergebnisse. Gustav Fischer Verlag, Jena, Germany.

27 Mai, Dieter, H., 2001. Die mittelmiozänen und obermiozänen Floren aus der Meuroer
28 und Raunoer Folge in der Lausitz. Teil II: Dicotyledonen. *Palaeontogr. Abteilung*
29 B, 35–174.

30 Mao, K. S., Hao, G., Liu, J. Q., Adams, R.P., Milne, R.I., 2010. Diversification and
31 biogeography of *Juniperus* (Cupressaceae): variable diversification rates and
32 multiple intercontinental dispersals. *New Phytol.* 188, 254–272.
33 <https://doi.org/10.1111/j.1469-8137.2010.03351.x>

34 Matzke, N.J., 2013. Probabilistic historical biogeography: new models for founder-
35 event speciation, imperfect detection, and fossils allow improved accuracy and
36 model-testing. *Front. Biogeogr.* 5. <https://doi.org/10.21425/f55419694>

37 McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., Brumfield, R.T., 2013.
38 Applications of next-generation sequencing to phylogeography and phylogenetics.
39 *Mol. Phylogenet. Evol.* 66, 526–538.
40 <https://doi.org/10.1016/j.ympev.2011.12.007>

- 1 Meng, Y., Wen, J., Nie, Z.L., Sun, H., Yang, Y.P., 2008. Phylogeny and biogeographic
2 diversification of *Maianthemum* (Ruscaceae: Polygonatae). *Mol. Phylogenet.*
3 *Evol.* 49, 424–434. <https://doi.org/10.1016/j.ympev.2008.07.017>
- 4 Menke B., 1976. Pliozäne und ältestquartäre Sporen- und Pollenflora von Schleswig-
5 Holstein. *Geol. Jahrb.* 32, 3–197.
- 6 Miao, Y. F., Herrmann, M., Wu, F. L., Yan, X., Yang, S. L., 2012. What controlled
7 Mid–Late Miocene long-term aridification in Central Asia? — Global cooling or
8 Tibetan Plateau uplift: A review. *Earth-Science Rev.* 112, 155–172.
9 <https://doi.org/10.1016/j.earscirev.2012.02.003>
- 10 Mutke, J., Barthlott, W., 2005. Patterns of vascular plant diversity at continental to
11 global scale. *Biol. Skr.* 55, 521–537.
- 12 Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000.
13 Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.
14 <https://doi.org/10.1038/35002501>
- 15 Palmer, J.D., Thompson, W.F., 1982. Chloroplast DNA rearrangements are more
16 frequent when a large inverted repeat sequence is lost. *Cell* 29, 537–550.
17 [https://doi.org/10.1016/0092-8674\(82\)90170-2](https://doi.org/10.1016/0092-8674(82)90170-2)
- 18 Parks, M., Cronn, R., Liston, A., 2009. Increasing phylogenetic resolution at low
19 taxonomic levels using massively parallel sequencing of chloroplast genomes.
20 *BMC Biol.* 7, 84. <https://doi.org/10.1186/1741-7007-7-84>
- 21 Polissar, P.J., Freeman, K.H., Rowley, D.B., McInerney, F.A., Currie, B.S., 2009.
22 Paleoaltimetry of the Tibetan Plateau from D/H ratios of lipid biomarkers. *Earth*
23 *Planet. Sci. Lett.* 287, 64–76. <https://doi.org/10.1016/j.epsl.2009.07.037>
- 24 Popescu, S.-M., 2002. Repetitive changes in Early Pliocene vegetation revealed by
25 high-resolution pollen analysis: revised cyclostratigraphy of southwestern
26 Romania. *Rev. Palaeobot. Palynol.* 120, 181–202. [https://doi.org/10.1016/S0034-6667\(01\)00142-7](https://doi.org/10.1016/S0034-6667(01)00142-7)
- 28 Pyron, R.A., 2014. Biogeographic analysis reveals ancient continental vicariance and
29 recent oceanic dispersal in amphibians. *Syst. Biol.* 63, 779–797.
30 <https://doi.org/10.1093/sysbio/syu042>
- 31 Qiu, Y.-X., Fu, C.-X., Comes, H.P., 2011. Plant molecular phylogeography in China
32 and adjacent regions: Tracing the genetic imprints of Quaternary climate and
33 environmental change in the world's most diverse temperate flora. *Mol.*
34 *Phylogenet. Evol.* 59, 225–244. <https://doi.org/10.1016/j.ympev.2011.01.012>
- 35 Raab-Straube, E. Von, 2003. Phylogenetic relationships in *Saussurea* (*Compositae*,
36 *Cardueae*) sensu lato, inferred from morphological, ITS and *trnL-trnF* sequence
37 data, with a synopsis of *Himalaiella* gen. nov., *Lipschitzziella* and *Frolovia*.
38 *Willdenowia* 33, 379–402. <https://doi.org/10.3372/wi.33.33214>
- 39 Raab-Straube, E. von, 2017. Taxonomic revision of *Saussurea* subgenus *Amphilaena*
40 (*Compositae*, *Cardueae*). *Botanic Garden and Botanical Museum Berlin, Berlin*.

- 1 Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior
2 summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–
3 904. <https://doi.org/10.1093/sysbio/syy032>
- 4 Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference
5 under mixed models. *Bioinformatics* 19, 1572–4.
- 6 Rowley, D.B., Garzione, C.N., 2007. Stable Isotope-Based Paleoaltimetry. *Annu. Rev.*
7 *Earth Planet. Sci.* 35, 463–508.
8 <https://doi.org/10.1146/annurev.earth.35.031306.140155>
- 9 Samuel, R., Turner, B., Duangjai, S., Munzinger, J., Paun, O., Barfuss, M.H.J., Chase,
10 M.W., 2019. Systematics and evolution of the Old World Ebenaceae, a review
11 with emphasis on the large genus *Diospyros* and its radiation in New Caledonia.
12 *Bot. J. Linn. Soc.* 99–114.
- 13 Schattner, P., Brooks, A.N., Lowe, T.M., 2005. The tRNAscan-SE, snoscan and
14 snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.*
15 33, W686–W689. <https://doi.org/10.1093/nar/gki366>
- 16 Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L., 2007. Comparison of whole
17 chloroplast genome sequences to choose noncoding regions for phylogenetic
18 studies in angiosperms: The Tortoise and the hare III. *Am. J. Bot.* 94, 275–288.
19 <https://doi.org/10.3732/ajb.94.3.275>
- 20 Shi, Z., Raab-Straube, E., 2011. Saussurea group. Pp. 42–149 in: Wu, Z.Y., Raven, P.H.
21 and Hong, D.Y. (eds.), *Flora of China*, vol. 20–21, Asteraceae. Science Press,
22 Beijing; Missouri Botanical Garden Press, St. Louis.
- 23 Smith, S.A., O’Meara, B.C., 2012. TreePL: Divergence time estimation using penalized
24 likelihood for large phylogenies. *Bioinformatics* 28, 2689–2690.
25 <https://doi.org/10.1093/bioinformatics/bts492>
- 26 Soltis, D.E., Burleigh, G., Barbazuk, W.B., Moore, M.J., Soltis, P.S., 2010. Advances
27 in the use of next-generation sequence data in plant systematics and evolution.
28 *Acta Hort.* 193–206. <https://doi.org/10.17660/ActaHortic.2010.859.23>
- 29 Spicer, R.A., Harris, N.B.W., Widdowson, M., Herman, A.B., Guo, S., Valdes, P.J.,
30 Wolfe, J.A., Kelley, S.P., 2003. Constant elevation of southern Tibet over the past
31 15 million years. *Nature* 421, 622–624. <https://doi.org/10.1038/nature01356>
- 32 Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-
33 analysis of large phylogenies. *Bioinformatics* 30, 1312–1313.
34 <https://doi.org/10.1093/bioinformatics/btu033>
- 35 Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the
36 RAxML web servers. *Syst. Biol.* 57, 758–771.
37 <https://doi.org/10.1080/10635150802429642>
- 38 Straub, S.C.K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R.C., Liston, A., 2012.
39 Navigating the tip of the genomic iceberg: Next-generation sequencing for plant
40 systematics. *Am. J. Bot.* 99, 349–364. <https://doi.org/10.3732/ajb.1100335>

- 1 Stubbs, R.L., Folk, R.A., Xiang, C.L., Soltis, D.E., Cellinese, N., 2018. Pseudo-parallel
2 patterns of disjunctions in an Arctic-alpine plant lineage. *Mol. Phylogenet. Evol.*
3 123, 88–100. <https://doi.org/10.1016/j.ympev.2018.02.016>
- 4 Sun, B., Wang, Y.-F., Li, C.-S., Yang, J., Li, J.-F., Li, Y.-L., Deng, T., Wang, S.-Q.,
5 Zhao, M., Spicer, R.A., Ferguson, D.K., Mehrotra, R.C., 2015. Early Miocene
6 elevation in northern Tibet estimated by palaeobotanical evidence. *Sci. Rep.* 5,
7 10379. <https://doi.org/10.1038/srep10379>
- 8 Sun, Y., Wang, A., Wan, D., Wang, Q., Liu, J., 2012. Rapid radiation of *Rheum*
9 (*Polygonaceae*) and parallel evolution of morphological traits. *Mol. Phylogenet.*
10 *Evol.* 63, 150–158. <https://doi.org/10.1016/j.ympev.2012.01.002>
- 11 Susanna, A., Galbany-Casals, M., Romaschenko, K., Barres, L., Martín, J., Garcia-
12 Jacas, N., 2011. Lessons from *Plectocephalus* (*Compositae*, *Cardueae*-
13 *Centaureinae*): ITS disorientation in annuals and Beringian dispersal as revealed
14 by molecular analyses. *Ann. Bot.* 108: 263–277.
- 15 Susanna, A., Garcia-Jacas, N., 2007. Tribe *Cardueae* Cass. In: Kadereit, J. K. and
16 Jeffrey C. (eds.) *The Families and Genera of Vascular Plants*, vol. 8. Springer-
17 Verlag, Berlin, Heidelberg.
- 18 Thorne, R.F., 1987. *Phytogeography: Floristic Regions of the World*. Science 236, 100–
19 100. <https://doi.org/10.1126/science.236.4797.100>
- 20 Tu, T. Y., Volis, S., Dillon, M.O., Sun, H., Wen, J., 2010. Dispersals of *Hyoscyameae*
21 and *Mandragoreae* (*Solanaceae*) from the New World to Eurasia in the early
22 Miocene and their biogeographic diversification within Eurasia. *Mol. Phylogenet.*
23 *Evol.* 57, 1226–1237. <https://doi.org/10.1016/j.ympev.2010.09.007>
- 24 Wagstaff, S.J., Breitwieser, I., 2004. Phylogeny and classification of *Brachyglottis*
25 (*Senecioneae*, *Asteraceae*): An example of a rapid species radiation in New
26 Zealand. *Syst. Bot.* 29, 1003–1010. <https://doi.org/10.1600/0363644042450991>
- 27 Wang, Y., Deng, T., Biasatti, D., 2006. Ancient diets indicate significant uplift of
28 southern Tibet after ca. 7 Ma. *Geology* 34, 309. <https://doi.org/10.1130/G22254.1>
- 29 Wang, Y., Li, X., Hao, G., Liu, J., 2004. Molecular phylogeny and biogeography of
30 *Androsace* (*Primulaceae*) and the convergent evolution of cushion morphology
31 *Acta Phytotaxon. Sin.* 42, 481–499].
- 32 Wang, Y. F., Wang, J. L., Wu, Y. Q., Du, G. Z., 2009. Floristic study on the genus
33 *Saussurea* in Gansu and floristic relations with its adjacent regions. *Guihaia* 29,
34 103–110.
- 35 Wang, Y. J., Liu, J. Q., 2004. A preliminary investigation on the phylogeny of
36 *Saussurea* (*Asteraceae*: *Cardueae*) based on chloroplast DNA *trnL-F* sequences.
37 *Acta Phytotaxon. Sin.* 42, 136–153.
- 38 Wang, Y.J., Susanna, A., Von Raab-strube, E., Milne, R., Liu, J.Q., 2009. Island-like
39 radiation of *Saussurea* (*Asteraceae*: *Cardueae*) triggered by uplifts of the Qinghai-
40 Tibetan Plateau. *Biol. J. Linn. Soc.* 97, 893–903.

8312.2009.01225.x

Wang, Y.J., von Raab-Straube, E., Susanna, A., Liu, J.Q., 2013. *Shangwua* (Compositae), a new genus from the Qinghai-Tibetan Plateau and Himalayas. *Taxon* 62, 984–996. <https://doi.org/10.12705/625.19>

Wen, J., Zhang, J.Q., Nie, Z.L., Zhong, Y., Sun, H., 2014. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Front. Genet.* 5. <https://doi.org/10.3389/fgene.2014.00004>

Wu, Z.Y., 1988. Hengduan Mountain flora and her significance. *J. Japanese Bot.* 63, 297–311.

Xia, X., 2018. DAMBE7: New and improved tools for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* 35, 1550–1552. <https://doi.org/10.1093/molbev/msy073>

Xie, L., Yang, Z.Y., Wen, J., Li, D.Z., Yi, T.S., 2014. Biogeographic history of *Pistacia* (Anacardiaceae), emphasizing the evolution of the Madrean-Tethyan and the eastern Asian-Tethyan disjunctions. *Mol. Phylogenet. Evol.* 77, 136–146. <https://doi.org/10.1016/j.ympev.2014.04.006>

Xu, Q., Ding, L., Zhang, L., Cai, FL., Lai, QZ., Yang, D., Jing, LZ., 2013. Paleogene high elevations in the Qiangtang Terrane, central Tibetan Plateau. *Earth Planet. Sci. Lett.* 362, 31–42. <https://doi.org/10.1016/j.epsl.2012.11.058>

Xu, Q., Ding, L., Zhang, L., Yang, D., Cai, F., Lai, Q., Liu, J., Shi, R., 2010. Stable isotopes of modern herbivore tooth enamel in the Tibetan Plateau: Implications for paleoelevation reconstructions. *Chinese Sci. Bull.* 55, 45–54. <https://doi.org/10.1007/s11434-009-0543-2>

Yamada, K.D., Tomii, K., Katoh, K., 2016. Application of the MAFFT sequence alignment program to large data-reexamination of the usefulness of chained guide trees. *Bioinformatics* 32, 3246–3251. <https://doi.org/10.1093/bioinformatics/btw412>

Yi, T. S., Wen, J., Golan-Goldhirsh, A., Parfitt, D.E., 2008. Phylogenetics and reticulate evolution in *Pistacia* (Anacardiaceae). *Am. J. Bot.* 95, 241–251. <https://doi.org/10.3732/ajb.95.2.241>

Yu, Y., Harris, A.J., He, X. J., 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Mol. Phylogenet. Evol.* 56, 848–850. <https://doi.org/10.1016/j.ympev.2010.04.011>

Yuan, Q., Bi, Y. C., Chen, Y. S., 2015. *Diplazoptilon* (Asteraceae) is merged with *Saussurea* based on evidence from morphology and molecular systematics. *Phytotaxa* 236, 53–61. <https://doi.org/10.11646/phytotaxa.236.1.4>

Zachos, J., Mark, P., Sloan, L., Thomas, E., Katharina, B., 2001. Trends, rhythms, and aberrations in Global Climate 65 Ma to present. *Science* 292, 686–693. <https://doi.org/10.1126/science.1059412>

Zerbino, D.R., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly

1 using de Bruijn graphs. Genome Res. 18, 821–829.
2 <https://doi.org/10.1101/gr.074492.107>

3 Zhang, D., Gao, F. L., Li, W.X., Jakovlić, I., Zou, H., Zhang, J., Wang, G.T., 2018.
4 PhyloSuite: an integrated and scalable desktop platform for streamlined molecular
5 sequence data management and evolutionary phylogenetics studies. bioRxiv
6 489088. <https://doi.org/10.1101/489088>

7 Zhang, J. W., Nie, Z. L., Sun, H., 2009. Cytological study on the genus *Syncalathium*
8 (Asteraceae-Lactuceae), an endemic taxon to alpine scree of the Sino-Himalayas.
9 J. Syst. Evol. 47, 226–230. <https://doi.org/10.1111/j.1759-6831.2009.00024.x>

10 Zhang, M.-L., Fritsch, P.W., 2010. Evolutionary response of *Caragana* (Fabaceae) to
11 Qinghai-Tibetan Plateau uplift and Asian interior aridification. Plant Syst. Evol.
12 288, 191–199. <https://doi.org/10.1007/s00606-010-0324-z>

13 Zhang, N., Erickson, D.L., Ramachandran, P., Ottesen, A.R., Timme, R.E., Funk, V.A.,
14 Luo, Y., Handy, S.M., 2017. An analysis of *Echinacea* chloroplast genomes:
15 Implications for future botanical identification. Sci. Rep. 7, 216.
16 <https://doi.org/10.1038/s41598-017-00321-6>

17 Zhang, X.S., 2007. Vegetation Map of the People's Republic of China (1:1,000,000).
18 The Geological Publishing House, Beijing.

- subg. *Amphileena*
- subg. *Eriocoryne*
- subg. *Theodorea*
- subg. *Saussurea* sect. *Saussurea*
- subg. *Saussurea* sect. *Rosulascences*
- subg. *Saussurea* sect. *Lagurostemon*
- subg. *Saussurea* sect. *Laguranthera*
- subg. *Saussurea* sect. *Gymnocline*
- subg. *Saussurea* sect. *Strictae*
- outgroup

Ice-free temperature °C

Distribution of Variable sites

Distribution of Proportion-variable-sites

Distribution of Parsimony-informative-sites

Distribution of Proportion-parsimony-informative-sites

- subg. *Amphilaena*
- subg. *Eriocoryne*
- subg. *Theodorea*
- subg. *Saussurea*, sect. *Saussurea*
- subg. *Saussurea*, sect. *Rosulascentes*
- subg. *Saussurea*, sect. *Lagrostemon*
- subg. *Saussurea*, sect. *Laguranthera*
- subg. *Saussurea*, sect. *Gymnocline*
- subg. *Saussurea*, sect. *Strictae*
- outgroup

Species names	Collection number	Localities		Rawreads number	Chloroplast length	Genebank number
<i>Arctium lappa</i>	M. Tang & C. Ren RW741	China,	Xizang, Gyirong	27064082	152708	MH671331
<i>Atractylodes lancea</i>					153099	KY977990
<i>A. macrocephala</i>					153265	MF034020
<i>Aucklandia costus</i>	Y. S. Chen 11-136	China,	Yunnan, Heqing	19599508	152592	MH926063
<i>Boloocephalus saussureoides</i>	Y. S. Chen 13-0385	China,	Xizang, Lhünzê	18535550	152499	MH926064
<i>Cirsium arvense</i>					152855	KY562583
<i>C. eriophorum</i>					152557	KY562584
<i>C. japonicum</i> var <i>maackii</i>					152545	MF034024
<i>C. japonicum</i> var <i>spinosissimum</i>					152550	MF034025
<i>C. vulgare</i>					152567	KY562585
<i>Carthamus tinctorius</i>					152938	KX822074
<i>Centaurea diffusa</i>					152559	KJ690264
<i>Dolomiaea souliei</i>	Kham Expedition 10-580	China,	Sichuan, Baiyu	20481246	152459	MH926065
<i>Hemistepha lyrata</i>	Q. Yuan 15-1	China,	Beijing	38446530	152454	MH926066
<i>Himalaiella deltoidea</i>	L. S. Xu 150177	China,	Hubei, Shennongjia	16756704	152579	MH926067
<i>Saussurea alaschanica</i>	Y. S. Chen 141000	China,	Inner Mongolia, Alashan	12166608	152562	MH926068
<i>S. thoroldii</i>	L. S. Xu 150365	China,	Qinghai, Delingha	28322332	152446	MH926174
<i>S. tsoongii</i>	L. S. Xu 150397	China,	Qinghai, Yushu	15831810	152567	MH926178
<i>S. alpina</i>	C. G. Alm 2822	Sweden,	Lappland, Jukkasjarvi	20316030	152471	MH926069
<i>S. amara</i>	Y. S. Chen 141565	China,	Xinjiang, Toli	12542792	152511	MH926070
<i>S. americana</i>	L. F. Hudson724	USA,	Alaska, Baranof Island	43119620	152471	MH926071
<i>S. amurensis</i>	L. S. Xu X140145	China,	Inner Mongolia, Ke Yi River	22175992	152460	MH926072
<i>S. andryaloides</i>	Y. S. Chen 13-2273	China,	Xinjiang, Taxkorgan	23346586	152451	MH926073
<i>S. apus</i>	L. S. Xu 150391	China,	Qinghai, Kunlun Mountains	25625432	152332	MH926074

<i>S. baicalensis</i>	Q. Yuan 15-3	China, Beijing	22312728	152577	MH926075
<i>S. baroniana</i>	Y. S. Chen 8134	China, Shaanxi, Mei Xian	19164618	152555	MH926076
<i>S. bartholomewii</i>	L. S. Xu 150448	China, Qinghai, Nangqên	32705520	152606	MH926077
<i>S. bhutanensis</i>	Y. S. Chen 13-2131	China, Xizang, Yadong	19220962	152589	MH926078
<i>S. brachycephala</i>	L. Yuan 151	Japan, Honshu, Iwate	31939466	152527	MH926079
<i>S. bracteata</i>	Y. S. Chen 13-1286	China, Xizang, Comai	20241738	152283	MH926080
<i>S. brunneopilosa</i>	Kham Expedition 10-1032	China, Qinghai, Yushu	14128286	152484	MH926081
<i>S. bullockii</i>	L. S. Xu XS180254	China, Fujian, Wuyishan	21193722	152491	MK966027
<i>S. cana</i>	Z. S. Zhang 119	China, Xinjiang, Hejing	12490722	152439	MH926082
<i>S. centiloba</i>	FLPH Sichuan Expedition 151472	China, Sichuan, Muli	45235918	152389	MH926083
<i>S. chabyoungsanica</i>	K-NU91468	Korea, Gangwon-do		152446	MH926084
<i>S. chinduensis</i>	Kham Expedition 10-931	China, Qinghai, Chinduo	25628392	152489	MH926085
<i>S. ciliaris</i>	Y. S. Chen 11-069	China, Sichuan, Yanyuan	25054276	152551	MH926086
<i>S. controversa</i>	M. N. Shurupova 10004	Russia, Republic of Khakassia, Efremkino	32854268	152518	MH926087
<i>S. coriolepis</i>	FLPH Sichuan Expedition 150855	China, Sichuan, Dayi	28040400	152600	MH926088
<i>S. davurica</i>	M. N. Shurupova 10005	Russia, Republic of Khakassia, The Lake Belyo	15235582	152213	MH926089
<i>S. delavayi</i>	Y. S. Chen 11-199	China, Yunnan, Dali	21038592	152254	MH926090
<i>S. depsangensis</i>	Y. S. Chen 13-1285	China, Xizang, Comai	26642626	152476	MH926091
<i>S. dolichopoda</i>	L. S. Xu XS150252	China, Guizhou, Tongren	18855270	152516	MK966028
<i>S. dulongjiangensis</i>	Jin et al. ST1223	China, Yunnan, Dulongjiang	37720138	152394	MH926092
<i>S. elegans</i>	Y. S. Chen 141592	China, Xinjiang, Yumin	19956614	152554	MH926093
<i>S. eriocephala</i>	FLPH Sichuan Expedition 151989	China, Sichuan, Muli	31638310	152434	MH926094
<i>S. eriostemon</i>	Y. S. Chen 404	China, Xizang,	20440152	152445	MH926095

		Gyirong			
<i>S. fuscipappa</i>	Y. S. Chen 11-191	China, Xizang, Zayü	25737604	152467	MH926096
<i>S. glabrescens</i>	Y. S. Chen & L. S. Xu 160237	China, Sichuan, Muli	20911908	152516	MH926098
<i>S. globosa</i>	Y. S. Chen 9627	China, Sichuan, Litang	17809962	152453	MH926099
<i>S. gossipiphora</i>	FLPH Tibet Expedition 12-0354	China, Xizang, Nyalam	33893076	152270	MH926100
<i>S. graminea</i>	Kham Expedition 10-1023	China, Xizang, Gonjo	21921982	152506	MH926101
<i>S. grosseserrata</i>	Y. S. Chen 11-107	China, Yunnan, Lijiang	13430914	152528	MH926102
<i>S. henryi</i>	L. S. Xu 150187	China, Hubei, Shennongjia	26803652	152582	MH926103
<i>S. hylophila</i>	X. T. Ma et al. PE6542	China, Xizang, Cona	23045082	152627	MH926104
<i>S. integrifolia</i>	Y. S. Chen & Z. H. Wang 9129	China, Sichuan, Xiaojin	19529100	152584	MH926105
<i>S. involucrata</i>	Y. S. Chen 141390	China, Xinjiang, Urumqi	23774034	152476	MH926106
<i>S. japonica</i>	Y. S. Chen 140531	China, Inner Mongolia, XilinGol	28198774	152545	MH926107
<i>S. kaschgarica</i>	Y. S. Chen 13-2304	China, Xinjiang, Wuqia	20136072	152483	MH926109
<i>S. katochaete</i>	Y. S. Chen 13-2210	China, Xizang, Rinbung	22384212	152475	MH926110
<i>S. kingii</i>	Y. S. Chen 9754	China, Xizang, Langkazi	33198726	152532	MH926111
<i>S. komaroviana</i>	L. S. Xu X140079	China, Jilin, Changbai Shan	25207628	152501	MH926112
<i>S. kuschakewiczii</i>	Y. S. Chen 13-2291	China, Xinjiang, Wuqia	15610466	152468	MH926113
<i>S. lanata</i>	FLPH Tibet Expedition 12-2144	China, Xizang, Gongbo'gyamda	27345504	152350	MH926114
<i>S. langpoensis</i>	X. T. Ma et al. PE6542	China, Xizang, Lampug	17287344	152206	MH926115
<i>S. laniceps</i>	Kham Expedition 10-3124	China, Yunnan, Shangri-la	19958614	152161	MH926116
<i>S. latifolia</i>	M. N. Shurupova 10003	Russia, Republic of Khakassia	22193102	152576	MH926117
<i>S. leucophylla</i>	Y. S. Chen 141408	China, Xinjiang, Hejing	41147882	152495	MH926118
<i>S. licentiana</i>	Y. S. Chen 8123	China, Shaanxi, Mei	28289270	152555	MH926119

		xian				
<i>S. longifolia</i>	Y. S. Chen 9633	China, Sichuan,	20119130	152489	MH926120	
		Daocheng				
<i>S. loriformis</i>	Y. S. Chen 9616	China, Sichuan,	20331288	152471	MH926121	
		Litang				
<i>S. malitiosa</i>	L. S. Xu 150376	China, Qinghai,	20241980	152499	MH926122	
		Delingha				
<i>S. merinoi</i>	Y. S. Chen & Z. H. Wang 9044	China, Yunnan, Qiaojia	19873122	152559	MH926123	
<i>S. mucronulata</i>	Z. S. Zhang ZZS160	China, Xinjiang, Akto	15337986	152526	MH926124	
<i>S. muliensis</i>	Y. S. Chen & L. S. Xu 160287	China, Sichuan, Muli	17874342	152584	MH926125	
<i>S. nigrescens</i>	Kham Expedition 10-1075	China, Qinghai, Yushu	29806346	152529	MH926126	
<i>S. nuda</i>	E. Lepage 23943	USA, Alaska, Nome River	37309652	152564	MH926127	
<i>S. obvallata</i>	Y. S. Chen 13-1796	China, Xizang, Yadong	40240324	152506	MH926128	
<i>S. odontolepis</i>	Y. S. Chen 140562	China, Inner Mongolia, Hexigten	19562950	152519	MH926129	
<i>S. orgaadayi</i>	Y. S. Chen 141668	China, Xinjiang, Qinghe	20564414	152555	MH926130	
<i>S. pachyneura</i>	Y. S. Chen & Z. H. Wang 9120	China, Sichuan, Xiaojin	21773144	152489	MH926131	
<i>S. pagriensis</i>	Y. S. Chen 13-2141	China, Xizang, Yadong	38195060	152377	MH926132	
<i>S. paleacea</i>	Kham Expedition 10-1857	China, Xizang, Zhag'yab	10576054	152396	MH926133	
<i>S. paucijuga</i>	L. S. Xu 150255	China, Shaanxi, Mei xian	14718014	152545	MH926134	
<i>S. peduncularis</i>	Kham Expedition 10-168	China, Yunnan, Shangri-la	25710410	152422	MH926135	
<i>S. petrovii</i>	L. S. Xu 150333	China, Qinghai, Ledu	15545822	152450	MH926136	
<i>S. picridifolia</i>	FLPH Tibet Expedition 12-1382	China, Xizang, Zayü	23161612	152364	MH926137	
<i>S. pinnatidentata</i>	L. S. Xu 150274	China, Shaanxi, Yulin	16353582	152524	MH926138	
<i>S. polylepis</i>	SKK044232	Korea, Hongdo Island		152488	MH926139	
<i>S. poochlamys</i>	Y. S. Chen 9638	China, Sichuan, Daocheng	29273472	152525	MH926140	
<i>S. populifolia</i>	Y. S. Chen 8116	China, Shaanxi, Mei	44776724	152522	MH926141	

			xian			
<i>S. pseudoalpina</i>	Y. S. Chen 141440		China, Xinjiang,	40240234	152560	MH926142
			Hejing			
<i>S. pseudomalitiosa</i>	Y. S. Chen 13-0019		China, Qinghai,	28494626	152565	MH926144
			Ulan			
<i>S. pseudorockii</i>	Y. S. Chen 9730		China, Yunnan,	19854992	152378	MH926146
			Gongshan			
<i>S. pseudosimpsoniana</i>	FLPH Tibet Expedition 12-1086		China, Xizang,	24988076	152379	MH926147
			Nang xian			
<i>S. pseudotridactyla</i>	Y. S. Chen 13-0994		China, Xizang,	11367174	152547	MH926148
			Cona,			
<i>S. pubifolia</i>	FLPH Tibet Expedition 12-1054		China, Xizang,	22104100	152663	MH926150
			Nang xian			
<i>S. pulchella</i>	L. S. Xu X140003		China, Jilin, Antu	37932620	152537	MH926151
<i>S. pulvinata</i>	L. S. Xu 150374		China, Qinghai,	15571344	152558	MH926152
			Delingha			
<i>S. romuleifolia</i>	Kham Expedition 10-2802		China, Yunnan,	37438674	152445	MH926153
			Shangri-la			
<i>S. runcinata</i>	Y. S. Chen 140956		China, Inner	23955782	152435	MH926154
			Mongolia, Alashan			
<i>S. salicifolia</i>	M. N. Shurupova 10001		Russia, Buryatia,	19721986	152461	MH926155
			Dushelan			
<i>S. salsa</i>	Y. S. Chen 141621		China, Xinjiang,	35797450	152542	MH926156
			Tuoli			
<i>S. semiamplexicaulia</i>	Kham Expedition 10-2103		China, Xizang,	22616094	152716	MH926157
			Markam			
<i>S. semifasciata</i>	FLPH Sichuan Expedition 152718		China, Sichuan,	38614298	152335	MH926158
			Daowu			
<i>S. semilyrata</i>	Kham Expedition 10-191		Chian, Yunnan,	26809274	152391	MH926159
			Shangri-la			
<i>S. simpsoniana</i>	Y. S. Chen 13-2279		Pakistan, Khunjerab	28202318	152024	MH926161
<i>S. sobarocephala</i>	Y. S. Chen 8121		China, Shaanxi, Mei	20743434	152514	MH926162
			xian			
<i>S. sobarocephaloides</i>	S. B. Lan 488		China, Yunnan,	40952896	152482	MH926163
			Dongchuan			
<i>S. sp. nov. 1</i> (subg. <i>Amphilaena</i>)	Y. S. Chen 13-0972		China, Xizang,	21058760	152476	MH926097
			Cona			
<i>S. sp. nov. 2</i> (subg. <i>Saussurea</i> sect. <i>Saussurea</i>)	FLPH Sichuan Expedition 151390		China, Sichuan,	20407410	152424	MH926108
			Muli			
<i>S. sp. nov. 3</i> (subg. <i>Eriocoryne</i>)	X. T. Ma et al. PE5956		China, Xizang,	17286872	152427	MH926143
			Lhünzê			
<i>S. sp. nov. 4</i> (subg. <i>Saussurea</i>)	Y. S. Chen 141562		China, Xinjiang,	24509782	152437	MH926145
			Toli			

<i>sect.</i>						
<i>Lagurostemon</i>)						
<i>S. sp. nov. 5</i>	Y. S. Chen 9232	China,	Sichuan,	20305938	152489	MH926160
<i>subg. Eriocoryne</i>		Sêrtar				
<i>S. sp. nov. 6</i>	FLPH Sichuan	China,	Sichuan,	20428438	152498	MH926188
<i>(subg. Saussurea</i>	Expedition 151871	Muli				
<i>sect. Saussurea)</i>						
<i>S. sp. nov. 7</i>	L. S. Xu 150319	China,	Gansu,	19616028	152468	MH926189
<i>(subg. Saussurea</i>		Tongren				
<i>sect. Saussurea)</i>						
<i>S. sp. nov. 8</i>	Kham Expedition	China,	Xizang,	17695498	152374	MH926190
<i>(subg. Saussurea</i>	10-1858	Zhag'yab				
<i>sect. Saussurea)</i>						
<i>S. sp. nov. 9</i>	X. T. Ma et al.	Cona,	Xizang,	19163714	152568	MH926182
<i>(subg. Amphilaena)</i>	PE6544	China				
<i>S. sp. nov. 10</i>	W. B. Ju AZH01323	China,	Sichuan,	27101974	152892	MH926149
<i>(subg. Eriocoryne)</i>		Chongzhou				
<i>S. spathulifolia</i>	Y. S. Chen 11-076	China,	Sichuan,	43050516	152443	MH926164
		Yanyuan				
<i>S. stella</i>	L. S. Xu 150424	China,	Qinghai,	13424530	152504	MH926165
		Yushu				
<i>S. stricta</i>	Y. S. Chen & L. S. Xu 160173	China,	Sichuan,	17428868	152660	MH926166
		Wenchuan				
<i>S. stubendorffii</i>	Y. S. Chen 141640	China,	Xinjiang,	19969772	152500	MH926167
		Burqin				
<i>S. subtriangulata</i>	L. S. Xu X140058	China, Jilin, Baishan		21668984	152583	MH926168
<i>S. subulata</i>	L. S. Xu 150386	China,	Qinghai,	18143232	152506	MH926169
		Kunlun Mountains				
<i>S. superba</i>	Y. S. Chen 13-0258	China,	Xizang,	41690464	152561	MH926170
		Lhünzê				
<i>S. sutchuenensis</i>	L. S. Xu 150166	China,	Hubei,	20536094	152494	MH926171
		Shennongjia				
<i>S. sutchuenensis</i>	L. S. Xu XS180253	China,	Guizhou,	18318524	152333	MK966029
		Tongren				
<i>S. tangutica</i>	Kham Expedition	China,	Qinghai,	45348880	152377	MH926172
	10-1027	Chinduo				
<i>S. thomsonii</i>	L. S. Xu 150388	China,	Qinghai,	24231684	152468	MH926173
		Kunlun Mountains				
<i>S. tianshuiensis</i>	L. S. Xu 150267	China,	Shaanxi,	33897420	152543	MH926175
		Feng xian				
<i>S. tomentosa</i>	L. S. Xu X140037	China, Jilin, Baishan		22542558	152519	MH926176
<i>S. tridactyla</i>	Y. S. Chen 13-0977	China,	Xizang,	20668738	152485	MH926177
		Cona				
<i>S. tunglingensis</i>	L. S. Xu X140156	China,	Inner	12155584	152510	MH926179

		Mongolia, Genhe			
<i>S. uliginosa</i>	Y. S. Chen 16277	China, Sichuan,	15459832	152507	MH926180
		Muli			
<i>S. uniflora</i>	Y. S. Chen 394	China, Xizang,	28892426	152345	MH926181
		Gyirong			
<i>S. veitchiana</i>	L. S. Xu 150181	China, Hubei,	30640136	152539	MH926183
		Shennongjia			
<i>S. velutina</i>	Y. S. Chen & L. S. Xu 160179	China, Sichuan,	17907502	152421	MH926184
		Wenchuan			
<i>S. wellbyi</i>	Y. S. Chen 13-0043	China, Qinghai, Hoh	44710296	152418	MH926185
		Xil			
<i>S. woodiana</i>	Y. S. Chen 9080	China, Sichuan,	44491492	152506	MH926186
		Mianning			
<i>S. xiaojinensis</i>	Y. S. Chen & L. S. Xu 160188	China, Sichuan,	23042166	152615	MH926187
		Xiaojin			
<i>Silybum marianum</i>				153202	KT267161

Alignment name	Variable sites	Proportion variable sites	Parsimony-informative sites	Proportion parsimony-informative sites
<i>trnH</i> -GUG	0	0	0	0
<i>trnH</i> -GUG to <i>psbA</i>	86	0.228	56	0.148
<i>psbA</i>	26	0.024	12	0.011
<i>psbA</i> to <i>trnK</i> -UUU				
exon 2	31	0.144	18	0.084
<i>trnK</i> -UUU exon 2	0	0	0	0
<i>trnK</i> -UUU exon 2 to				
matK	37	0.133	19	0.068
matK	168	0.111	93	0.061
matK to <i>trnK</i> -UUU				
exon 1	73	0.11	28	0.042
<i>trnK</i> -UUU exon 1	0	0	0	0
<i>trnK</i> -UUU exon 1 to				
<i>rps16</i> exon 2	124	0.158	55	0.07
<i>rps16</i> exon 2	10	0.052	3	0.016
<i>rps16</i> exon 2 to <i>rps16</i>				
exon 1	101	0.12	48	0.057
<i>rps16</i> exon 1	0	0	0	0
<i>rps16</i> exon 1 to <i>trnQ</i> -				
UUG	177	0.19	88	0.094
<i>trnQ</i> -UUG	4	0.056	4	0.056
<i>trnQ</i> -UUG to <i>psbK</i>	45	0.129	18	0.052
<i>psbK</i>	6	0.033	4	0.022
<i>psbK</i> to <i>psbI</i>	60	0.149	28	0.069
<i>psbI</i>	4	0.036	1	0.009
<i>psbI</i> to <i>trnS</i> -GCU	35	0.257	21	0.154
<i>trnS</i> -GCU	1	0.011	1	0.011
<i>trnS</i> -GCU to <i>trnC</i> -				
GCA	205	0.265	72	0.093
<i>trnC</i> -GCA	5	0.062	4	0.049
<i>trnC</i> -GCA to <i>petN</i>	159	0.283	70	0.125
<i>petN</i>	4	0.044	2	0.022
<i>petN</i> to <i>psbM</i>	90	0.196	39	0.085
<i>psbM</i>	8	0.076	3	0.029
<i>psbM</i> to <i>trnD</i> -GUC	90	0.14	37	0.058
<i>trnD</i> -GUC	1	0.014	0	0
<i>trnD</i> -GUC to <i>trnY</i> -				
GUA	17	0.159	10	0.093
<i>trnY</i> -GUA	0	0	0	0
<i>trnY</i> -GUA to <i>trnE</i> -				
UUC	17	0.327	14	0.269
<i>trnE</i> -UUC	10	0.139	8	0.111

<i>trnE</i> -UUC to <i>rpoB</i>	309	0.344	236	0.263
<i>rpoB</i>	160	0.05	60	0.019
<i>rpoB</i> to <i>rpoC1</i> exon 1	6	0.231	6	0.231
<i>rpoC1</i> exon 1	27	0.063	10	0.023
<i>rpoC1</i> exon 1 to				
<i>rpoC1</i> exon 2	45	0.062	23	0.032
<i>rpoC1</i> exon 2	90	0.055	28	0.017
<i>rpoC1</i> exon 2 to				
<i>rpoC2</i>	14	0.137	6	0.059
<i>rpoC2</i>	278	0.067	119	0.029
<i>rpoC2</i> to <i>rps2</i>	23	0.093	9	0.036
<i>rps2</i>	30	0.042	14	0.02
<i>rps2</i> to <i>atpI</i>	10	0.047	6	0.028
<i>atpI</i>	33	0.044	10	0.013
<i>atpI</i> to <i>atpH</i>	151	0.135	60	0.054
<i>atpH</i>	4	0.016	3	0.012
<i>atpH</i> to <i>atpF</i> exon 1	56	0.153	21	0.058
<i>atpF</i>	5	0.035	1	0.007
<i>atpF</i> exon 1 to <i>atpF</i>				
exon 2	78	0.11	27	0.038
<i>atpF</i> exon 2	19	0.046	9	0.022
<i>atpF</i> exon 2 to <i>atpA</i>	4	0.062	4	0.062
<i>atpA</i>	67	0.044	31	0.02
<i>atpA</i> to <i>trnR</i> -UCU	20	0.159	4	0.032
<i>trnR</i> -UCU	1	0.014	0	0
<i>trnR</i> -UCU to <i>trnG</i> -				
UCC exon 2	73	0.264	32	0.116
<i>trnG</i> -UCC exon 2	1	0.021	1	0.021
<i>trnG</i> -UCC exon 2 to				
<i>trnG</i> -UCC exon 1	52	0.074	19	0.027
<i>trnG</i> -UCC exon 1	0	0	0	0
<i>trnG</i> -UCC exon 1 to				
<i>trnT</i> -GGU	25	0.156	17	0.106
<i>trnT</i> -GGU	17	0.236	6	0.083
<i>trnT</i> -GGU to <i>psbD</i>	136	0.113	64	0.053
<i>psbD</i>	39	0.037	25	0.024
<i>psbD</i> to <i>trnS</i> -UGA	100	0.062	56	0.035
<i>trnS</i> -UGA	1	0.012	1	0.012
<i>trnS</i> -UGA to <i>psbZ</i>	29	0.086	15	0.045
<i>psbZ</i>	6	0.032	0	0
<i>psbZ</i> to <i>trnfM</i> -CAU	50	0.088	20	0.035
<i>trnfM</i> -CAU	1	0.014	0	0
<i>trnfM</i> -CAU to <i>rps14</i>	16	0.105	8	0.053
<i>rps14</i>	9	0.03	5	0.017

<i>rps14</i> to <i>psaB</i>	7	0.053	3	0.023
<i>psaB</i>	62	0.028	28	0.013
<i>psaB</i> to <i>psaA</i>	0	0	0	0
<i>psaA</i>	62	0.028	29	0.013
<i>psaA</i> to <i>ycf3</i> exon 3	74	0.101	25	0.034
<i>ycf3</i> exon 3	6	0.039	3	0.02
<i>ycf3</i> exon 3 to <i>ycf3</i> exon 2	54	0.073	27	0.036
<i>ycf3</i> exon 2	5	0.022	1	0.004
<i>ycf3</i> exon 2 to <i>ycf3</i> exon 1	85	0.122	47	0.068
<i>ycf3</i> exon 1	4	0.032	0	0
<i>ycf3</i> exon 1 to <i>trnS</i> - GGA	160	0.174	78	0.085
<i>trnS</i> -GGA	0	0	0	0
<i>trnS</i> -GGA to <i>rps4</i>	29	0.095	14	0.046
<i>rps4</i>	36	0.059	16	0.026
<i>rps4</i> to <i>trnT</i> -UGU	45	0.13	15	0.043
<i>trnT</i> -UGU	1	0.014	0	0
<i>trnT</i> -UGU to <i>trnL</i> - UAA exon 1	114	0.205	55	0.099
<i>trnL</i> -UAA exon 1	1	0.027	0	0
<i>trnL</i> -UAA exon 1 to <i>trnL</i> -UAA exon 2	32	0.073	14	0.032
<i>trnL</i> -UAA exon 2	3	0.06	0	0
<i>trnL</i> -UAA exon 2 to <i>trnF</i> -GAA	60	0.164	30	0.082
<i>trnF</i> -GAA	2	0.029	0	0
<i>trnF</i> -GAA to <i>ndhJ</i>	65	0.16	39	0.096
<i>ndhJ</i>	11	0.023	2	0.004
<i>ndhJ</i> to <i>ndhK</i>	9	0.088	5	0.049
<i>ndhK</i>	34	0.05	14	0.021
<i>ndhK</i> to <i>ndhC</i>	4	0.089	2	0.044
<i>ndhC</i>	14	0.039	8	0.022
<i>ndhC</i> to <i>trnV</i> -UAC exon 2	185	0.164	80	0.071
<i>trnV</i> -UAC exon 2	1	0.027	1	0.027
<i>trnV</i> -UAC exon 2 to <i>trnV</i> -UAC exon 1	33	0.058	21	0.037
<i>trnV</i> -UAC exon 1	0	0	0	0
<i>trnV</i> -UAC exon 1 to <i>trnM</i> -CAU	14	0.08	5	0.029
<i>trnM</i> -CAU	2	0.027	0	0
<i>trnM</i> -CAU to <i>atpE</i>	28	0.139	16	0.079

<i>atpE</i>	13	0.032	5	0.012
<i>atpB</i>	92	0.062	61	0.041
<i>atpB</i> to <i>rbcL</i>	85	0.112	40	0.053
<i>rbcL</i>	79	0.055	48	0.034
<i>rbcL</i> to <i>accD</i>	80	0.14	25	0.044
<i>accD</i>	108	0.077	55	0.039
<i>accD</i> to <i>psaI</i>	135	0.202	55	0.082
<i>psaI</i>	4	0.036	4	0.036
<i>psaI</i> to <i>ycf4</i>	43	0.106	19	0.047
<i>ycf4</i>	30	0.054	13	0.023
<i>ycf4</i> to <i>cemA</i>	134	0.158	73	0.086
<i>cemA</i>	107	0.155	18	0.026
<i>cemA</i> to <i>petA</i>	31	0.137	11	0.049
<i>petA</i>	52	0.054	21	0.022
<i>petA</i> to <i>psbJ</i>	111	0.142	38	0.049
<i>psbJ</i>	1	0.008	1	0.008
<i>psbJ</i> to <i>psbL</i>	10	0.068	2	0.014
<i>psbL</i>	2	0.025	1	0.012
<i>psbL</i> to <i>psbF</i>	4	0.069	2	0.034
<i>psbF</i>	4	0.033	3	0.025
<i>psbE</i>	5	0.02	1	0.004
<i>psbE</i> to <i>petL</i>	224	0.179	92	0.073
<i>petL</i>	5	0.052	1	0.01
<i>petL</i> to <i>petG</i>	43	0.267	10	0.062
<i>petG</i>	7	0.061	0	0
<i>petG</i> to <i>trnW-CCA</i>	17	0.145	11	0.094
<i>trnW-CCA</i>	0	0	0	0
<i>trnW-CCA</i> to <i>trnP-UGG</i>				
<i>trnP-UGG</i>	25	0.152	13	0.079
<i>trnP-UGG</i>	1	0.014	0	0
<i>trnP-UGG</i> to <i>psaJ</i>	35	0.114	20	0.065
<i>psaJ</i>	6	0.047	2	0.016
<i>psaJ</i> to <i>rpl33</i>	61	0.143	33	0.077
<i>rpl33</i>	20	0.1	11	0.055
<i>rpl33</i> to <i>rps18</i>	24	0.135	9	0.051
<i>rps18</i>	10	0.033	5	0.016
<i>rps18</i> to <i>rpl20</i>	37	0.151	17	0.069
<i>rpl22</i> to <i>rps19</i>	7	0.113	2	0.032
<i>rpl23</i>	2	0.007	1	0.004
<i>clpP</i> exon 3	12	0.053	6	0.026
<i>clpP</i> exon 3 to <i>clpP</i> exon 2				
<i>clpP</i> exon 2	51	0.081	22	0.035
<i>clpP</i> exon 2	22	0.075	14	0.048

<hr/>				
<i>clpP</i> exon 2 to <i>clpP</i>				
exon 1	70	0.088	36	0.045
<i>clpP</i> exon 1	3	0.042	3	0.042
<i>clpP</i> exon 1 to <i>psbB</i>	40	0.089	20	0.045
<i>psbB</i>	78	0.051	41	0.027
<i>psbB</i> to <i>psbT</i>	32	0.16	15	0.075
<i>psbT</i>	8	0.078	6	0.059
<i>psbT</i> to <i>psbN</i>	6	0.082	2	0.027
<i>psbN</i>	3	0.023	2	0.015
<i>psbN</i> to <i>psbH</i>	4	0.039	3	0.029
<i>psbH</i>	15	0.068	6	0.027
<i>psbH</i> to <i>petB</i>	116	0.111	52	0.05
<i>petB</i>	15	0.035	6	0.014
<i>petB</i> to <i>petD</i>	79	0.094	29	0.034
<i>petD</i>	15	0.029	5	0.01
<i>petD</i> to <i>rpoA</i>	30	0.154	13	0.067
<i>rpoA</i>	71	0.071	23	0.023
<i>rpoA</i> to <i>rps11</i>	7	0.092	2	0.026
<i>rps11</i>	17	0.041	9	0.022
<i>rps11</i> to <i>rpl36</i>	12	0.114	8	0.076
<i>rpl36</i>	3	0.026	0	0
<i>rpl36</i> to <i>infA</i>	20	0.174	7	0.061
<i>infA</i>	12	0.051	5	0.021
<i>infA</i> to <i>rps8</i>	21	0.174	12	0.099
<i>rps8</i>	26	0.064	11	0.027
<i>rps8</i> to <i>rpl14</i>	24	0.133	6	0.033
<i>rpl14</i>	12	0.033	5	0.014
<i>rpl14</i> to <i>rpl16</i>	20	0.172	6	0.052
<i>rpl16</i>	18	0.05	6	0.017
<i>rpl16</i> to <i>rps3</i>	138	0.114	58	0.048
<i>rps3</i>	43	0.065	26	0.04
<i>rpl23</i> to <i>trnI</i> -CAU	3	0.018	0	0
<i>rpl2</i> exon 1	3	0.008	0	0
<i>rps19</i>	14	0.093	8	0.053
<i>rps19</i> to <i>rpl2</i> exon 2	37	0.203	8	0.044
<i>rpl20</i> to <i>clpP</i> exon 3	93	0.092	45	0.045
<i>rpl22</i>	45	0.095	18	0.038
<i>rpl20</i>	20	0.052	13	0.034
<i>rpl2</i> exon 2	3	0.007	2	0.005
<i>rpl2</i> exon 2 to <i>rpl2</i>				
exon 1	6	0.009	6	0.009
<i>trnI</i> -CAU	0	0	0	0
<i>trnI</i> -CAU to <i>ycf2</i>	1	0.009	1	0.009
<i>ycf2</i>	175	0.026	108	0.016
<hr/>				

<i>ycf2</i> to <i>trnL</i> -CAA	3	0.007	1	0.002
<i>trnL</i> -CAA	1	0.012	1	0.012
<i>trnL</i> -CAA to <i>ndhB</i>				
exon 2	16	0.028	6	0.011
<i>ndhB</i> exon 2	9	0.012	5	0.007
<i>ndhB</i> exon 2 to <i>ndhB</i>				
exon 1	9	0.014	2	0.003
<i>ndhB</i> exon 1	8	0.01	1	0.001
<i>ndhB</i> exon 1 to <i>rps7</i>	14	0.048	9	0.031
<i>rps7</i>	4	0.009	3	0.006
<i>rps7</i> to <i>ycf15</i>	42	0.023	13	0.007
<i>ycf15</i>	6	0.031	3	0.016
<i>ycf15</i> to <i>trnV</i> -GAC	26	0.038	8	0.012
<i>trnV</i> -GAC	1	0.014	0	0
<i>trnV</i> -GAC to <i>trnI</i> -				
GAU exon 1	110	0.055	19	0.009
<i>trnI</i> -GAU exon 1	0	0	0	0
<i>trnI</i> -GAU exon 1 to				
<i>trnI</i> -GAU exon 2	11	0.012	5	0.005
<i>trnI</i> -GAU exon 2	0	0	0	0
<i>trnI</i> -GAU exon 2 to				
<i>trnA</i> -UGC exon 1	1	0.016	0	0
<i>trnA</i> -UGC exon 1	0	0	0	0
<i>trnA</i> -UGC exon 1 to				
<i>trnA</i> -UGC exon 2	8	0.01	6	0.007
<i>trnA</i> -UGC exon 2	0	0	0	0
<i>trnA</i> -UGC exon 2 to				
<i>trnR</i> -ACG	58	0.015	19	0.005
<i>trnR</i> -ACG	0	0	0	0
<i>trnR</i> -ACG to <i>trnN</i> -				
GUU	25	0.054	7	0.015
<i>trnN</i> -GUU	0	0	0	0
<i>trnN</i> -GUU to <i>ycf1</i>	9	0.028	2	0.006
<i>ycf1</i>	43	0.075	31	0.054
<i>ndhF</i>	237	0.106	109	0.049
<i>ndhF</i> to <i>rpl32</i>	244	0.233	88	0.084
<i>rpl32</i>	15	0.091	4	0.024
<i>rpl32</i> to <i>trnL</i> -UAG	203	0.247	93	0.113
<i>trnL</i> -UAG	0	0	0	0
<i>trnL</i> -UAG to <i>ccsA</i>	24	0.169	14	0.099
<i>ccsA</i>	88	0.091	36	0.037
<i>ccsA</i> to <i>ndhD</i>	44	0.18	18	0.073
<i>ndhD</i>	85	0.061	37	0.027
<i>ndhD</i> to <i>psaC</i>	23	0.096	12	0.05

<i>psaC</i>	7	0.028	3	0.012
<i>psaC</i> to <i>ndhE</i>	37	0.157	16	0.068
<i>ndhE</i>	15	0.049	6	0.02
<i>ndhE</i> to <i>ndhG</i>	13	0.06	7	0.032
<i>ndhG</i>	33	0.062	13	0.024
<i>ndhG</i> to <i>ndhI</i>	62	0.182	32	0.094
<i>ndhI</i>	41	0.082	18	0.036
<i>ndhI</i> to <i>ndhA</i> exon 2	9	0.115	8	0.103
<i>ndhA</i> exon 2	24	0.044	12	0.022
<i>ndhA</i> exon 2 to <i>ndhA</i> exon 1	176	0.166	66	0.062
<i>ndhA</i> exon 1	36	0.065	18	0.033
<i>ndhH</i>	78	0.066	41	0.035
<i>ndhH</i> to <i>rps15</i>	14	0.154	6	0.066
<i>rps15</i>	25	0.093	12	0.044
<i>rps15</i> to end	1113	0.22	697	0.137

1 **Explanations of appendices**

2

3 Appendix Table 1. Sample information.

4

5 Appendix Table 2. Number of Variable sites, proportion of variable sites, number
6 parsim-informative sites and proportion of parsim-informative sites of each region.

7 The order is according to the gene order of chloroplast genomes.

8

9 Appendix Fig.1 Distribution of Variable sites, proportion of variable sites, parsim-
10 informative sites and proportion of parsim-informative sites of each region.

11

12 Appendix Fig.2 BI tree inferred from all coding region dataset. Numbers above and
13 below branches are Bayesian posterior probability (PP) from BI analyze and bootstrap
14 values (BS) from ML, respectively. PP<0.95 and BS<75 is marked with red color.

15

16 Appendix Fig.3 ML tree from 30 most informative regions, bootstrap values (BS)
17 marked. BS<75 is marked with red color.

18