Molecular Phylogenetics and Evolution 50 (2009) 547–559

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogeny of *Caragana* (Fabaceae) based on DNA sequence data from *rbcL*, *trnS–trnG*, and ITS

Mingli Zhang [a,b], Peter W. Fritsch [c,*], Boni C. Cruz [c]

[a] State Key Lab of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, PR China
[b] Laboratory of Evolution and Biodiversity Conservation in Arid Regions, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, Xinjiang, PR China
[c] Department of Botany, California Academy of Sciences, 55 Music Concourse Drive, Golden Gate Park, San Francisco, CA 94118-4503, USA

ARTICLE INFO

Article history:
Received 16 July 2008
Revised 22 November 2008
Accepted 1 December 2008
Available online 8 December 2008

Keywords:
Caragana
Fabaceae
ITS
Phylogeny
rbcL
trnS–trnG

ABSTRACT

Phylogenetic relationships of 48 species of *Caragana* (Fabaceae: tribe Hedysareae) and one representative each of *Astragalus*, *Calophaca*, *Halimodendron*, and *Hedysarum* are estimated from DNA sequences of the *rbcL* gene, *trnS–trnG* intron and spacer, and ITS region. At least one representative of all five sections and 12 series within *Caragana* are included. Analyses yielded strongly supported clades corresponding to sections *Caragana*, *Bracteolatae*, and *Frutescentes*. The species of section *Jubatae* are distributed among three strongly supported clades, i.e., one with the species of section *Bracteolatae*, another with two species of section *Spinosae*, and a third as sister to section *Frutescentes*. All but the last of these six clades are corroborated by at least one unambiguously traced morphological character. The placement of the other four species of section *Spinosae* are not well supported and lack unambiguous morphological synapomorphies, and the samples of *Calophaca* and *Halimodendron* nest within *Caragana* with weak support.

1. Introduction

Caragana Fabr. (Fabaceae: Papilionoideae; Polhill, 1981; Lock, 2005) comprises about 100 species distributed in northern Eurasia, from the Black Sea to southeastern Siberia, south to eastern and southwestern China, Nepal, Afghanistan, and Turkmenistan. It commonly occurs in cold arid regions, such as the Qinghai–Xizang (Tibet) Plateau, but also is found in forested areas of eastern Asia, especially in northern China. *Caragana* species often form the dominant component of the natural vegetation in cold-temperate dry and arid scrublands, montane meadows, and deserts (Wu, 1980; Zhang et al., 2002). Because of their adaptation to arid conditions, many species of *Caragana* are widely used as ground covers to control soil erosion in dry areas. They are also used as windbreaks, living fences, shade trees, and ornamentals. *Caragana arborescens* is frequently cultivated in North America.

Based on morphological similarity, Polhill (1981) placed *Caragana* in subtribe Astragalinae of tribe Galegeae. A phylogenetic estimate of the "temperate herbaceous clade" of Papilionoideae based on DNA sequence data from the internal transcribed spacer (ITS) region of nuclear ribosomal DNA yielded a strongly supported clade (bootstrap value (bt) = 95) comprising the two sampled species of *Caragana*, *Calophaca* Fisch. ex DC., and *Halimodendron* Fisch ex DC. to the exclusion of the other members of the subtribe (i.e., *Alhagi* Gagnebin, *Astragalus* L., *Biserrula* L., *Chesneya* Lindl. ex Endl., *Gueldenstaedtia* Fisch., and *Oxytropis* DC.) as well as *Hedysarum* L. and *Onobrychis* Mill. (Sanderson and Wojciechowski, 1996). A supertree analysis based on sequence data from various genic regions also recovered this clade, with a clade comprising *Alhagi*, *Hedysarum* L., and *Onobrychis* Mill. as its sister (Wojciechowski et al., 2000). An analysis based on *matK* sequences with a subset of these taxa yielded a clade comprising *Alhagi*, *Caragana*, *Hedysarum*, and *Onobrychis* with strong support (bt = 85–92; Bayesian posterior probability (pP) = 1.00), with *Caragana* as sister to a clade comprising the remaining taxa (bt, pP < 50; Wojciechowski et al., 2004). A subsequent analysis combining *matK* and ITS data recovered the same clade (bt < 50; pP = 1.00) with *Caragana* as sister (bt = 100; pP = 1.00; Wojciechowski, 2005). Consequently, Lock (2005) transferred these four genera plus *Calophaca* and *Halimodendron* to tribe Hedysareae.

Although there is strong support for a clade comprising *Caragana*, *Calophaca*, and *Halimodendron*, molecular phylogenetic studies are thus far inconclusive regarding relationships among these genera. An analysis based on ITS sequences, limited to four species of the group, resulted in a clade of *Cal. tianschanica* (B. Fedtsch.) Boriss. and *Car. frutex* (bt = 73) as sister to *Car. arborescens* (bt = 79); this larger clade was in turn sister to *Halimodendron* (Sanderson and Wojciechowski, 1996). A supertree approach based on various genic regions with five samples of the clade resulted in the same

* Corresponding author. Fax: +1 415 379 5737.
E-mail address: pfritsch@calacademy.org (P.W. Fritsch).

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.12.001

topology, with *Cal. wolgarica* Fisch. as sister to the rest (Wojciechowski et al., 2000). These analyses indicate that *Caragana* may not be monophyletic.

Komarov (1908) published the first monograph of *Caragana*, in which eight series were delimited (*Caragana* (=*Altaganae*, not validly published; McNeill et al., 2006: Art. 22.2), *Bracteolatae* Kom., *Erinacanthae* Kom., *Frutescentes* Kom., *Jubatae* Kom., *Occidentales* Kom., *Pygmaeae* Kom., and *Spinosae* Kom.). Series *Caragana* was subdivided into subseries *Caragana* (=*Arborescentes*, not validly published; McNeill et al., 2006), *Microphyllae* Kom., and *Stipitatae* Kom. Subsequently, two authors modified this classification in the context of regional floras (e.g., Pojarkova, 1945, *Flora of the USSR*; Liu, 1993, *Flora of the People's Republic of China*), and others proposed revisions of the entire genus (Sanchir, 1980, 1999; Gorbunova 1984; Zhao, 1993). Although some of these contributed to the classification of new species, all (including that of Komorov) are of limited utility because they were based only on overall similarity rather than modern concepts of monophyly.

Moore (1968) reported chromosome counts for 17 species of *Caragana* and combined these data with two characters traditionally important in infrageneric classification (leaf rachis development (deciduous versus persistent) and foliage condition (pinnate versus digitate)) to provide the first phylogenetic estimate of the genus. No explanation was provided in the study as to how the tree was derived, although the positions of the subdivisions were imprecisely plotted on a graph with the axes corresponding to the two characters with the most derived states farthest from the origin. Moore found that most species sampled are diploid ($2n = 16$), and proposed that the species of series *Chamlagu* Pojark. (segregated from ser. *Frutescentes* by Pojarkova, 1945), with triploid and hexaploid chromosome complements, were derived by allopolyploid speciation from hybridization between species in series *Microphyllae* (diploid) and *Frutescentes* (diploid and tetraploid).

Subsequently, chromosome counts have been reported for 11 additional species (Zhou et al., 2002), and the pollen morphology of 34 species has also been described (Zhang et al., 1996). Through the availability of these new data, Zhang (1997) conducted a phylogenetic analysis based on gross morphological data, chromosome numbers, and pollen characters within a modern phylogenetic framework. Based on these results, Zhang (1997) revised the infrageneric classification of the genus above the species level, recognizing five sections (*Caragana*, *Bracteolatae*, *Frutescentes*, *Jubatae*, and *Spinosae*), each with two or three series (*Caragana*, *Microphyllae*; *Ambiguae*, *Bracteolatae*; *Chamlagu*, *Frutescentes*, *Pygmaeae*; *Jubatae*, *Leucospinae*; and *Acanthophyllae*, *Dasyphyllae*, *Spinosae*, respectively; Table 1).

Recently, the phylogeny of *Caragana* was estimated with sequence data from the ITS region of nuclear ribosomal DNA, and the *trnL–trnF* intron/spacer and *trnS*^{GCU}–*trnG*^{UUC}spacer regions of chloroplast DNA, with 20 Chinese species of the genus. This study provided only limited insight into the phylogeny of *Caragana* because (1) only one-fifth of the genus was sampled, (2) the samples are restricted to Chinese species, and (3) the only other genus sampled, *Sophora japonica* L. (tribe Sophoreae; used as the outgroup), is only distantly related to *Caragana* relative to genera in the Hedysareae. Here we assess previous classifications of *Caragana* based on morphology with a phylogenetic estimate from additional ITS DNA sequences, and the *trnS*^{GCU}–*trnG*^{UUC}–*trnG*^{UUC} spacer/intron region (''*trnS–trnG* region") and *rbcL* gene of chloroplast DNA. We improve upon the study of Hou et al. (2008) by including more species and the *trnG*^{UUC}–*trnG*^{UUC} intron portion of the *trnS–trnG* region. We also trace the evolution of morphological characters considered important in the infraspecific classification of the genus over a phylogenetic estimate from the analysis of the combined 3-gene data to assess the relative utility of these characters for clade diagnosis.

2. Materials and methods

2.1. Taxon sampling

Forty-eight species of *Caragana* were sampled for this study. This represents ca. 48% of the total number of species in the genus and covers all series and sections sensu Zhang (1997; Table 1). The sample includes two accessions of *C. microphylla* (Mongolia and Berlin Botanical Garden, Germany, originally from northern China). One species each of the genera *Halimodendron* (*H. halodendron* (Pall.) Voss.) and *Calophaca* (*C. soongorica* Kar. & Kir.) were included as members of the outgroup. Because of the uncertainty regarding the monophyly of *Caragana* with respect to *Calophaca* and possibly *Halimodendron* (Sanderson and Wojciechowski, 1996; Wojciechowski et al., 2000), *Hedysarum alpinum* L. was included as an additional member of the outgroup. This genus is placed in the sister clade of *Caragana*, *Halimodendron*, and *Calophaca* (Wojciechowski et al., 2000, 2004; Wojciechowski, 2005). A member of the Astragalean clade (*Astragalus*) was also included as a more distant member of the outgroup on the basis of data from Wojciechowski et al. (2004). *Astragalus coluteocarpus* Boiss. was used for the ITS and *trnS–trnG* regions, *A. sparsus* Decne. for *rbcL*, and these data were combined into a single terminal for analysis.

2.2. DNA sequencing

The isolation of total DNA followed the protocol in Wang et al. (2004). The polymerase chain reaction (PCR) was performed with standard methods (Dieffenbach and Dveksler, 1995) and either BIOLASE (Bioline USA, Randolph, MA, USA), iTaq (Bio-Rad Laboratories, Inc., Hercules, CA, USA), or HotStart-IT (USB Corporation, Cleveland, OH, USA) as the DNA polymerase. The PCR products were purified by using Exo I/SAP (USB Corporation). Cycle sequencing was performed with the ABI Prism BigDye Terminators v 3.1 Cycle Sequencing Reaction Kit (Applied Biosystems, Foster City, CA, USA) by using 1/8-scale reaction mixtures in a model 9600 PCR System thermal cycler (Perkin-Elmer, Boston, MA, USA) or MyCycler thermal cycler (Bio-Rad Laboratories, Inc.), and sequences were determined with an ABI Prism 3100 Genetic Analyzer (Applied Biosystems).

Amplification and sequencing of the ITS region employed primers from Swensen et al. (1998). The *trnS–trnG* spacer and *trnG* intron were amplified separately. Amplification and sequencing of the *trnS–trnG* region employed primers trnS^{GCU}, 3′trnG^{UUC}, 5′trnG2G, and 5′trnG2S from Shaw et al. (2005). The *rbcL* gene was amplified with primers rbcL 12–3′ (5′-CTC GGA GCT CCT TTT AGT AAA AGA TTG GGC CGA G-3′) and rbcL 1B-5′ (5′-ATG TCA CCA CAA ACA GAA ACT AAA GCA AGT-3′; Olmstead et al., 1992). Sequencing of the *rbcL* gene employed primers Z-234F (5′-CGT TAT AAA GGA CGA TGC TAC CAC ATC GA-3′), Z-234R-A (5′-TCG ATG TGG TAG CAT CGT CCT TTA TAA CG-3′), Z-674F-A (5′-TTT ATA AAG CAC AGG CTG AAA CAG GTG AAA TC-3′), and Z-674R-A (5′-GAT TTC ACC TGT TTC AGC CTG TGC TTT ATA AA-3′), which are modifications of primers from G. Zurawski (DNAX research Institute, Palo Alto, CA, USA), and rbcL 1351R (5′-CTT CAC AAG CAG CAG CTA GTT CAG GAC TCC-3′) from the laboratory of M.W. Chase. Forward and reverse sequences were edited by using the computer program Sequencher (4.1.2 and 4.2; Gene Codes Corp., Ann Arbor, MI, USA). Gaps introduced into the alignment were treated as missing data. All sequences have been deposited in GenBank (Table 1).

2.3. Phylogenetic analysis

Sequences from the ITS and *trnS–trnG* regions were manually aligned; no alignment was required for the *rbcL* region. Data set

Table 1
Voucher information for the 48 species of *Caragana* (49 samples) and four outgroups.

Taxon	Voucher	Source	GenBank Accession Nos. (ITS, *rbcL*, *trnS–trnG*)
Sect. *Caragana*[a]			
Ser. *Caragana*			
C. arborescens Lam.	*M.L. Zhang 00-201* (PE)	Altai, Xinjiang, China	FJ537262, FJ537211, FJ537164
C. boisii C.K. Schneid.	*M.L. Zhang & Y. Kang 00-121* (PE)	Lixian, Sichuan, China	FJ537259, FJ537208, FJ537161
C. prainii C.K. Schneid.	*D. Podlech 16678* (MSB)	Kunar, Afghanistan	FJ537255, FJ537205, FJ537157
C. purdomii Rehder	*C.Y. Chang* et al. *2004059* (WUG)	Yan'an, Shaanxi, China	FJ537261, FJ537710, FJ537163
C. soongorica Grubov	*M.L. Zhang* 00-256 (PE)	Cultivated, Urumqi Botanical Garden, Xinjiang, China	FJ537257, FJ537207, FJ537159
C. stipitata Kom.	*Y. Kang 00-55* (PE)	Huashan (Qingling), Shaanxi, China	FJ537260, FJ537209, FJ537162
C. turkestanica Kom.	*M.L. Zhang 00-101* (PE)	Cultivated, Bergius Botanical Garden, Stockholm, Sweden	FJ537256, FJ537206, FJ537158
C. zahlbruckneri C.K. Schneid.	*S.Y. He 18765* (PE)	Zhangjiakou, Hebei, China	FJ537258, –, FJ537160
Ser. *Microphyllae* (Kom.) Pojark.			
C. bungei Ledeb.	*M.L. Zhang* et al. *99-225* (PE)	Bajanchongor, Mongolia	FJ537267, FJ537216, FJ537169
C. korshinskii Kom.	*M.L. Zhang 00-149* (PE)	Cultivated, Turfan Botanical Garden, Xinjiang, China	FJ537266, FJ537215, FJ537168
C. microphylla Lam.1	*M.L. Zhang* et al. *99-214* (PE)	Lhongcheng, Mongolia	FJ537264, FJ537213, FJ537166
C. microphylla Lam. 2	*M.L. Zhang 177-99-74-80* (PE)	Cultivated, Berlin Botanical Garden, Germany; originally from northern China	FJ537265, FJ537214, FJ537167
C. pekinensis Kom.	*M.L. Zhang 99-56* (PE)	Xiangshan, Beijing, China	FJ537263, FJ537212, FJ537165
Sect. *Bracteolatae* (Kom.) M.L. Zhang			
Ser. *Bracteolatae* Kom.			
C. bicolor Kom.	*M.L. Zhang & Y. Kang Y 99-178* (PE)	Markang, Sichuan, China	FJ537246, FJ537197, FJ537147
C. brevispina Benth.	*M.L. Zhang 281-05-8414/101* (PE)	Cultivated, Berlin Botanical Garden, Germany (originally from Kashmir)	FJ537248, FJ537200, FJ537150
C. franchetiana Kom.	*M.L. Zhang & S.Z. Zhang 94-178* (WUG)	Gongbujiangda, Xizang, China	–, FJ537198, FJ537148
C. sukiensis C.K. Schneid.	*S.G. Miehe & K. Kock s.n.* (NHM)	Donkardzong, Nepal	FJ537247, FJ537199, FJ537149
Ser. *Ambiguae* Sanchir			
C. ambigua Stocks	*R.P. Steward 28001* (K)	Baluchistan, Pakistan	FJ537249, –, FJ537151
C. conferta Benth. ex Baker	*J.F. Duthie 12192* (NHM)	Astor-Gudhui, Kashmir	FJ537250, –, FJ537152
Sect. *Jubatae* (Kom.) Y.Z. Zhao			
Ser. *Jubatae* Kom.			
C. jubata (Pall.) Poir.	*M.L. Zhang 00279* (PE)	Zhaosu (Tianshan), Xinjiang, China	FJ537242, FJ537194, FJ537143
C. pleiophylla (Regel) Pojark.	*M.L. Zhang 10-146* (PE)	Tekes, Xinjiang, China	FJ537253, FJ537203, FJ537155
C. roborovskyi Kom.	*M.L. Zhang 00-88* (PE)	Uhai, Nei Mongol, China	FJ537254, FJ537204, FJ537156
C. tangutica Maxim.	*Q.L. Ho* et al. *2499* (NHM)	Yushu, Qinghai, China	FJ537278, FJ537227, FJ537180
Ser. *Leucospinae* Y.Z. Zhao			
C. changduensis Y.X. Liou	*Z.C. Ni* et al. *1069* (PE)	Chayü, Xizang, China	FJ537243, –, FJ537144
C. gerardiana Benth.	*S.G. Miehe & K. Kock K 01-032-03* (NHM)	Western Nepal	FJ537245, FJ537196, FJ537146
C. tibetica (Maxim. ex C.K. Schneid.) Kom.	*M.L. Zhang 00-89* (PE)	Uhai, Nei Mongol, China	FJ537244, FJ537195, FJ537145
Sect. *Frutescentes* (Kom.) Sanchir			
Ser. *Frutescentes* Kom.			
C. camilli–schneideri Kom.	*C.Y. Chang* et al. *2004334* (WUG)	Yumin, Xinjiang, China	FJ537283, FJ537232, FJ537184
C. frutex (L.) K. Koch	*M.L. Zhang 177-97-74-80* (PE)	Cultivated, Berlin Botanical Garden, Germany	FJ537285, FJ537234, FJ537186
C. kirghisorum Pojark.	*C.Y. Chang* et al. *2004219* (WUG)	Khorgos, Xinjiang, China	FJ537280, FJ537229, FJ537181
C. laeta Kom.	*M.L. Zhang 177-98-74-80* (PE)	Cultivated, Berlin Botanical Garden, Germany	FJ537281, FJ537230, FJ537182
C. opulens Kom.	*M.L. Zhang & Y. Kang 99-123* (PE)	Daofu, Sichuan, China	FJ537282, FJ537231, FJ537183
C. polourensis Franch.	*B. Bartholomew* et al. *9417* (CAS)	Minfeng (Kunlun), Xinjiang, China	FJ537279, FJ537228, –
Ser. *Chamlagu* Pojark.			
C. rosea Turcz. ex Maxim.	*M.L. Zhang 99-45* (PE)	Beihuashan, Beijing, China	FJ537272, FJ537221, FJ537174
C. sinica (Buc'hoz) Rehder	*M.L. Zhang 99-49* (PE)	Xiangshan, Beijing, China	FJ537284, FJ537233, FJ537185
C. ussuriensis (Regel) Pojark.	*M.L. Zhang 00-113* (PE)	Cultivated, Uppsala Botanical Garden, Sweden	FJ537273, FJ537222, FJ537175
Ser. *Pygmaeae* Kom.			
C. aurantiaca Koehne	*M.L. Zhang 00-156* (PE)	Cultivated, Turfan Botanical Garden, Xinjiang, China	FJ537270, FJ537219, FJ537172
C. brevifolia Kom.	*Q.L. Ho* et al. *2498* (NHM)	Yushu, Qinghai, China	FJ537268, FJ537217, FJ537170
C. chinghaiensis Y.X. Liou	*Q.L. Ho* et al. 93 (CAS)	Tongde, Qinghai, China	FJ537269, FJ537218, FJ537171
C. gobica Sanchir	*M.L. Zhang* et al. *99-304* (PE)	Gobi-Altai, Mongolia	FJ537277, FJ537226, FJ537179
C. leucophloea Pojark.	*M.L. Zhang* et al. *99-218* (PE)	Daxinchileng, Mongolia	FJ537275, FJ537224, FJ537177
C. pygmaea (L.) DC.	*M.L. Zhang 00-187* (PE)	Jinghe, Xinjiang, China	FJ537276, FJ537225, FJ537178
C. stenophylla Pojark.	*M.L. Zhang* 00-78 (PE)	Hangjinqi, Nei Mongol, China	FJ537274, FJ537223, FJ537176
C. versicolor Benth.	*S. Miehe 99-62-06* (NHM)	Upper Dolpo, Nepal	FJ537271, FJ537220, FJ537173
Sect. *Spinosae* (Kom.) Y.Z. Zhao			
Ser. *Spinosae* Kom.			
C. bongardiana (Fisch. & C.A. Mey.) Pojark.	*M.L. Zhang 00215* (PE)	Jimunai, Xinjiang, China	FJ537251, FJ537201, FJ537153
C. hololeuca Bunge ex Kom.	*M.L. Zhang 00-153* (PE)	Cultivated, Turfan Botanical Garden, Xinjiang, China	FJ537240, FJ537192, FJ537141
C. spinosa (L.) Hornem.	*C.Y. Chang* et al. *2004503* (WUG)	Qinghe, Xinjiang, China	FJ537241, FJ537193, FJ537142
C. tragacanthoides (Pall.) Poir.	*C.Y. Chang* et al. *2004404* (WUG)	Hebukesaier, Xinjiang, China	FJ537252, FJ537202, FJ537154
Ser. *Acanthophyllae* Pojark.			
C. acanthophylla Kom.	*M.L. Zhang 00-154* (PE)	Cultivated, Turfan Botanical Garden, Xinjiang, China	FJ537238, FJ537191, FJ537139
Ser. *Dasyphyllae* Pojark.			
C. dasyphylla Pojark.	*Xinjiang Expedition Team 472* (WUG)	Kuche, Xinjiang, China	FJ537239, –, FJ537140

(*continued on next page*)

Table 1 (*continued*)

Taxon	Voucher	Source	GenBank Accession Nos. (ITS, *rbcL*, *trnS–trnG*)
Outgroup			
Astragalus coluteocarpus Boiss.	*Qinghai–Xizang Expedition Team 76-8083* (PE)	Zada, Ali, Xizang, China	FJ537286, –, FJ537187
Astregalus sparsus Decne.			–, –, Z95550
Calophaca soongorica Kar. & Kir.	*E.E. Pyoahobeq & L.A. Kpamapehko 5-14-1984* (PE)	Semiipalatinskaya, Tajikstan	FJ537288, FJ537237, FJ537189
Halimodendron halodendron (Pall.) Voss	*M.L. Zhang 00-279* (PE)	Cultivated, Urumqi Botanical Garden, Xinjiang, China	FJ537289, FJ537237, FJ537190
Hedysarum alpinum L.	*M. Riewe 182* (CAS)	Northwest Territories, Canada	FJ537287, FJ537235, FJ537188

[a] The classification of *Caragana* follows Zhang (1997).

congruence was determined with incongruence length difference (ILD) tests (Farris et al., 1994). This test was implemented in the computer program PAUP* version 4.0b10 (Swofford, 2002) as described in Wang et al. (2004). The data sets compared were *trnS–trnG* versus *rbcL* and nuclear data versus chloroplast data. From the results of the ILD tests, three data sets were employed in analyses: ITS, cpDNA (*trnS–trnG* region + *rbcL*), and combined 3-gene. DNA from the following species failed to amplify: *Caragana franchetiana* for ITS; *C. palourensis* for the *trnS–trnG* region; and *C. ambigua*, *C. changduensis*, *C. conferta*, *C. dasyphylla*, and *C. zahlbruckneri* for *rbcL*. These species were excluded from the specific data set for which genic region data were missing, and from the combined 3-gene data set.

Phylogenetic analyses employed maximum parsimony (MP), maximum likelihood (ML), and Bayesian inference (BI). The MP analyses were conducted with the heuristic search option in PAUP*. Searches were conducted over 100 random-taxon-addition replicates with tree bisection-reconnection branch-swapping, steepest descent, and MulTrees in effect. All characters and states were weighted equally and unordered. All trees from the replicates were swapped to completion, all shortest trees were saved, and a strict consensus tree was computed. Relative support for individual clades was estimated with the parsimony bootstrap (bt) method (Felsenstein, 1985). One thousand pseudoreplicates were performed with uninformative characters excluded. Ten random-taxon-addition heuristic searches for each pseudoreplicate were performed and all minimum-length trees were saved per search.

The ML analyses were conducted with the heuristic search option in PAUP. The most complex model (GTR + I + Γ) was employed, in accordance with the recommendations of Huelsenbeck and Rannala (2004). Base frequencies and model parameters were estimated from the data, and four iterations were completed. Initial parameters were estimated from a neighbor-joining tree.

Bayesian analyses were conducted with MrBayes version 3.0b4 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) by using uniform prior probabilities and estimating base frequencies and the parameters for the GTR + I + Γ model as above. We ran four chains of the Markov chain Monte Carlo by beginning with a random tree and sampling one tree every 100 generations for 5,000,000 generations. The first 50,000 generations of the chain were used as "burn in" after stationarity was reached, and the phylogenetic estimate was based on trees sampled after generation 50,000. To estimate the posterior probability (pP) of recovered branches, 50% majority-rule consensus trees were created.

The evolution of 18 parsimony-informative morphological characters that have been used in previous infrageneric classifications of *Caragana* or to distinguish *Caragana* from *Calophaca* and *Halimodendron* were optimized onto the combined 3-gene ML tree by using Fitch parsimony with MacClade 4.0 (Maddison and Maddison, 2000; Tables 2 and 3). For completeness, taxa that were excluded from the combined 3-gene analysis because of missing genic regions were added to the combined tree at positions inferred from the analysis of individual data sets. *Caragana gobica* was excluded from this analysis because the available morphological data were too incomplete for this species. To achieve complete resolution in the tree (a requirement of the analysis) polytomies were arbitrarily resolved. The morphological characters and their states are based on the data set of Zhang (1997), as modified by species description data from Liu et al. (in press). To assess the distribution of characters and their states in the optimization of characters onto the combined tree, the following goodness-of-fit character indices were calculated with PAUP*: consistency index

Table 2
The 18 morphological characters of *Caragana* and outgroup taxa.

1. Leaves paripinnate (0); leaves imparipinnate (1)
2. Leaflet arrangement pinnate (0); leaflet arrangement digitate (1)
3. Leaflet pair number 5–10 (0); leaflet pair number (2–)3–4 (1); leaflet pair number 2 (2)
4. Leaf rachis + petiole deciduous (0); leaf rachis of long branches persistent and sclerotic, those of short branches deciduous (1); leaf rachis of both long and short branches persistent and sclerotic (2)
5. Leaf rachis + petiole length > 2 cm (0); leaf rachis length 1–2 cm (1); leaf rachis length < 1 cm (2)
6. Leaflet length/width < 2 (0); leaflet length/width 2–3 (1); leaflet length/width > 3 (2). The coding reflects leaflet shape: (0) more or less orbicular; (1) elliptic to narrowly ovate, (2) lanceolate to linear
7. Inflorescence fasciculate, geminate, or 1-flowered (0); inflorescence racemose (1)
8. Pedicel articulated at or above middle (0); pedicel articulated below middle or articulation absent (1)
9. Calyx shape campanulate (0); calyx shape campanulate-tubular or tubular (1)
10. Calyx base not gibbous (0); calyx base gibbous (1)
11. Calyx teeth length/tube length < 1/3 (0); calyx teeth length/tube length ⩾ 1/3 (1)
12. Corolla post-anthesis yellow or orange (0); corolla post-anthesis red or amaranth (1)
13. Standard broadly rounded or broadly obovate (0); standard broadly lanceolate or narrowly obovate (1)
14. Wing auricle length/claw length ⩽ 1/3 (0); wing auricle length/claw length > 1/3 (1)
15. Pollen exine ornamentation perforate (0); pollen exine reticulate (1)
16. Pod distinctly dehiscent (0); fruit not distinctly dehiscent (1)
17. Pod neck longer than calyx tube (0); pod neck shorter than calyx tube (1)
18. Pod inner wall glabrous (0); pod inner wall pubescent (1)

Characters and their states are modified from Zhang (1997).

Table 3
Morphological character matrix of *Caragana* and outgroup taxa.

	0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1
	1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8
Caragana acanthophylla	0 0 1 1 0 0 0 A 1 0 0 0 0 0 ? 0 1 0
C. ambigua	0 0 1 1 1 0 0 1 0 0 1 0 0 0 ? 0 1 1
C. arborescens	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
C. aurantiaca	0 1 2 1 2 2 0 1 0 0 0 0 0 1 1 0 1 0
C. bicolor	0 0 0 1 0 0 0 1 0 0 1 0 0 1 1 0 1 1
C. boisii	0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0
C. bongardiana	0 0 1 2 0 1 0 1 1 0 0 0 0 1 ? 0 1 1
C. brevifolia	0 1 2 1 2 1 0 1 1 0 0 0 0 0 1 0 1 0
C. brevispina	0 0 0 1 0 0 0 0 0 0 1 0 0 1 ? 0 1 1
C. bungei	0 0 1 0 0 0 0 0 1 0 0 0 0 0 ? 0 1 0
C. camilli–schneideri	0 1 2 1 2 0 0 0 1 1 0 0 0 0 1 0 1 1
C. changduensis	0 0 0 2 0 1 0 1 1 0 1 0 0 0 ? 0 1 0
C. chinghaiensis	0 1 2 2 2 1 0 1 0 0 1 0 0 1 ? 0 1 0
C. conferta	0 0 1 1 0 1 0 1 0 0 1 0 ? 1 ? 0 1 1
C. dasyphylla	0 A 2 1 B 1 0 1 1 0 0 0 0 1 0 0 1 0
C. franchetiana	0 0 0 1 0 0 0 1 0 0 1 0 0 1 1 0 1 1
C. frutex	0 1 2 1 2 1 0 0 1 1 0 0 0 1 0 0 1 0
C. gerardiana	0 0 1 2 0 1 0 1 1 0 1 0 0 0 0 0 1 1
C. hololeuca	0 0 2 1 1 1 0 1 1 0 0 0 0 1 ? 0 1 1
C. jubata	0 0 0 2 0 1 0 1 1 0 1 1 0 1 1 0 1 0
C. kirghisorum	0 1 2 1 2 1 0 0 1 1 0 0 1 0 ? 0 1 0
C. korshinskii	0 0 0 0 0 0 0 0 1 0 0 0 0 0 ? 0 1 0
C. laeta	0 1 2 1 B 0 0 0 1 1 0 0 1 0 ? 0 1 0
C. leucophloea	0 1 2 1 2 2 0 A 0 0 0 0 0 1 0 0 1 0
C. microphylla	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0
C. opulens	0 1 2 1 2 1 0 0 1 1 0 0 0 0 1 0 1 0
C. pekinensis	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1
C. pleiophylla	0 0 0 2 0 1 0 1 1 0 1 0 1 1 ? 0 1 0
C. polourensis	0 1 2 1 2 1 0 A 1 1 0 0 0 0 1 0 1 0
C. prainii	0 0 A 0 A 0 0 0 0 0 0 0 0 0 ? 0 1 0
C. purdomii	0 0 0 0 0 0 0 0 1 0 0 0 0 0 ? 0 0 0
C. pygmaea	0 1 2 1 2 2 0 A 1 0 0 0 0 0 0 0 1 1
C. roborovskyi	0 0 A 2 A 0 0 1 1 0 1 0 1 1 ? 0 1 0
C. rosea	0 1 2 1 2 1 0 A 1 1 0 1 1 0 1 0 1 0
C. sinica	0 A 2 1 B 1 0 0 0 0 0 0 0 1 0 1 0 0 0
C. soongorica	0 0 1 0 0 0 0 0 0 0 0 0 0 0 ? 0 1 0
C. spinosa	0 0 1 1 0 2 0 1 1 0 0 0 0 0 ? 0 1 0
C. stenophylla	0 1 2 1 2 2 0 1 1 0 0 0 0 1 1 0 1 0
C. stipitata	0 0 0 0 0 0 0 0 0 0 0 0 0 0 ? 0 0 0
C. sukiensis	0 0 0 1 0 0 0 1 0 0 1 0 0 0 ? 0 0 1
C. tangutica	0 0 1 2 0 A 0 1 1 0 0 0 0 1 1 0 1 0
C. tibetica	0 0 1 2 0 2 0 1 1 0 0 0 1 0 1 0 1 1
C. tragacanthoides	0 0 1 1 A 2 0 1 1 0 0 0 1 1 ? 0 1 1
C. turkestanica	0 0 A 0 0 0 0 0 0 0 0 0 0 0 ? 0 1 0
C. ussuriensis	0 A 2 1 B 1 0 0 0 A 0 1 1 0 ? 0 1 0
C. versicolor	0 1 2 1 2 2 0 1 1 0 0 0 0 0 1 0 1 0
C. zahlbruckneri	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0
Calophaca soongorica	1 0 A 0 0 0 1 0 1 1 1 0 0 0 ? 0 0 0
Halimodendron halodendron	0 0 1 1 0 1 1 1 0 1 0 1 0 1 ? 1 0 0
Astragalus coluteocarpus	1 0 0 0 0 0 1 0 0 0 1 1 0 0 ? 1 0 0
Hedysarum alpinum	1 0 0 0 0 1 1 0 0 0 0 1 0 0 ? 1 0 0

Characters and their states are modified from Zhang (1997). A, 0 and 1; B, 1 and 2; ?, missing.

(CI), retention index (RI), and homoplasy index (HI). Calculations were performed by interpreting taxa with multiple states as polymorphic.

3. Results

The data partitions *trnS–trnG* region versus *rbcL* and ITS versus cpDNA were not significantly incongruent on the basis of the ILD tests (both $P = 0.056$). Therefore we combined all cpDNA data into a single data set for a cpDNA analysis, and combined all data (ITS + cpDNA) into a single 3-gene data set for a combined analysis. Data set and tree statistics for the various MP analyses are summarized in Table 4.

The MP, ML, and BI analyses all resulted in similar trees in each of the data sets; no differences were supported by bt values > 80 or pP values > 0.95. There are often differences among the trees from the different analyses involving non-resolution (polytomies), but for brevity we will describe only instances of incongruence.

3.1. ITS analysis

The ML analysis resulted in five equally optimal trees (score = 2678.5175) that differ only in the relative placement of *Caragana bungei*, *C. korshinskii*, and *C. microphylla* 1. In the strict consensus of these trees, both *Calophaca* and *Halimodendron* nest within *Caragana* (bt < 50; pP = 0.98), the former as sister to *C. hololeuca* (bt < 50; pP < 0.50), the latter as part of a polytomy with two other clades (bt < 50; pP = 0.71; Fig. 1). Section *Caragana* is monophyletic (bt = 67; pP = 1.00), as is section *Bracteolatae* (bt < 50; pP = 1.00), and sect. *Frutescentes* is monophyletic except for the inclusion of *C. tangutica* (bt = 72; pP = 0.84). Neither sections *Jubatae* nor *Spinosae* are monophyletic. The species of sect. *Jubatae* are distributed among four clades (all bt < 50; pP $\leqslant$ 0.87), as are the species of sect. *Spinosae* (all bt < 50; pP $\leqslant$ 0.91). None of the series that are represented by more than one terminal are unequivocally monophyletic.

There are no instances of incongruence between the ML and BI trees (not shown). The only incongruence between the MP and ML trees is that in the MP tree *Caragana kirghisorum* is sister to *C. laeta* + *C. polourensis*, and *C. versicolor* is sister to a clade of *C. frutex*, *C. kirghisorum*, *C. laeta*, *C. opulens*, *C. polourensis*, and *C. sinica* (not shown).

3.2. cpDNA analysis

The ML analysis resulted in one optimal tree (score = 7191.7971). *Halimodendron* is sister to *Caragana* + *Calophaca* (bt < 50; pP < 50), and *Calophaca* forms a polytomy with the three major clades of *Caragana* species (Fig. 2). Sections *Caragana*, *Bracteolatae*, and *Frutescentes* are all monophyletic (bt = 91, <50, =68; pP = 1.00, 0.67, 0.97, respectively). The species of sect. *Jubatae* are distributed among three clades, with *C. tangutica* as the first-diverging lineage (bt = 0.68; pP = 0.97) of a clade that also includes sect. *Frutescentes* (bt = 90; pP = 0.90); *C. jubata* and a clade of *C. tibetica* + *C. gerardiana* group with sect. *Bracteolatae* (bt = 83;

Table 4
Data set and tree statistics from separate maximum parsimony analyses of ITS, *trnS–trnG*, *rbcL*, *trnS–trnG* + *rbcL*, and combined 3-gene for *Caragana*.

Data set statistics					Tree statistics				
Genic region	Aligned length (bp)	Number (%) variable characters	Number (%) parsimony-informative characters	% Missing data cells	Number of shortest trees	Length	CI[a]	RI[b]	Number of nodes in strict consensus/ max.no. (%)
ITS	643	197 (30.6)	104 (16.2)	0.06	3120	331	0.731	0.809	37/51(= 72.5)
trnS–trnG	1435	290 (20.2)	112 (7.8)	7.8	45,896	189	0.735	0.862	23/51(= 45.1)
rbcL	1428	119 (8.3)	40 (2.8)	0.5	13,069	79	0.595	0.861	19/47(= 40.4)
trnS–trnG + *rbcL*	2863	406 (14.2)	146 (5.1)	2.4	912	270	0.667	0.842	28/46(= 60.9)
Combined 3-gene	3506	600 (17.1)	235 (6.7)	1.5	275	481	0.617	0.807	34/52(= 65.4)

[a] Consistency index.
[b] Retention index.

pP = 0.97) as successive sister lineages to the latter with weak support; and *C. pleiophylla* and *C. roborovskyi* group with *C. bongariana* and *C. tragacanthoides* of sect. *Spinosae* (bt = 97; pP = 1.00). As for the other species of sect. *Spinosae*, *C. hololeuca* and *C. spinosa* each form part of a trichotomy with a clade comprising members of sects. *Jubatae* and *Bracteolatae* (bt < 50; pP < 0.50), and *C. acanthophylla* is the first-diverging lineage (bt = 50; pP = 0.96) of a clade that also includes the clades of sect. *Caragana* and *C. bongardi-*

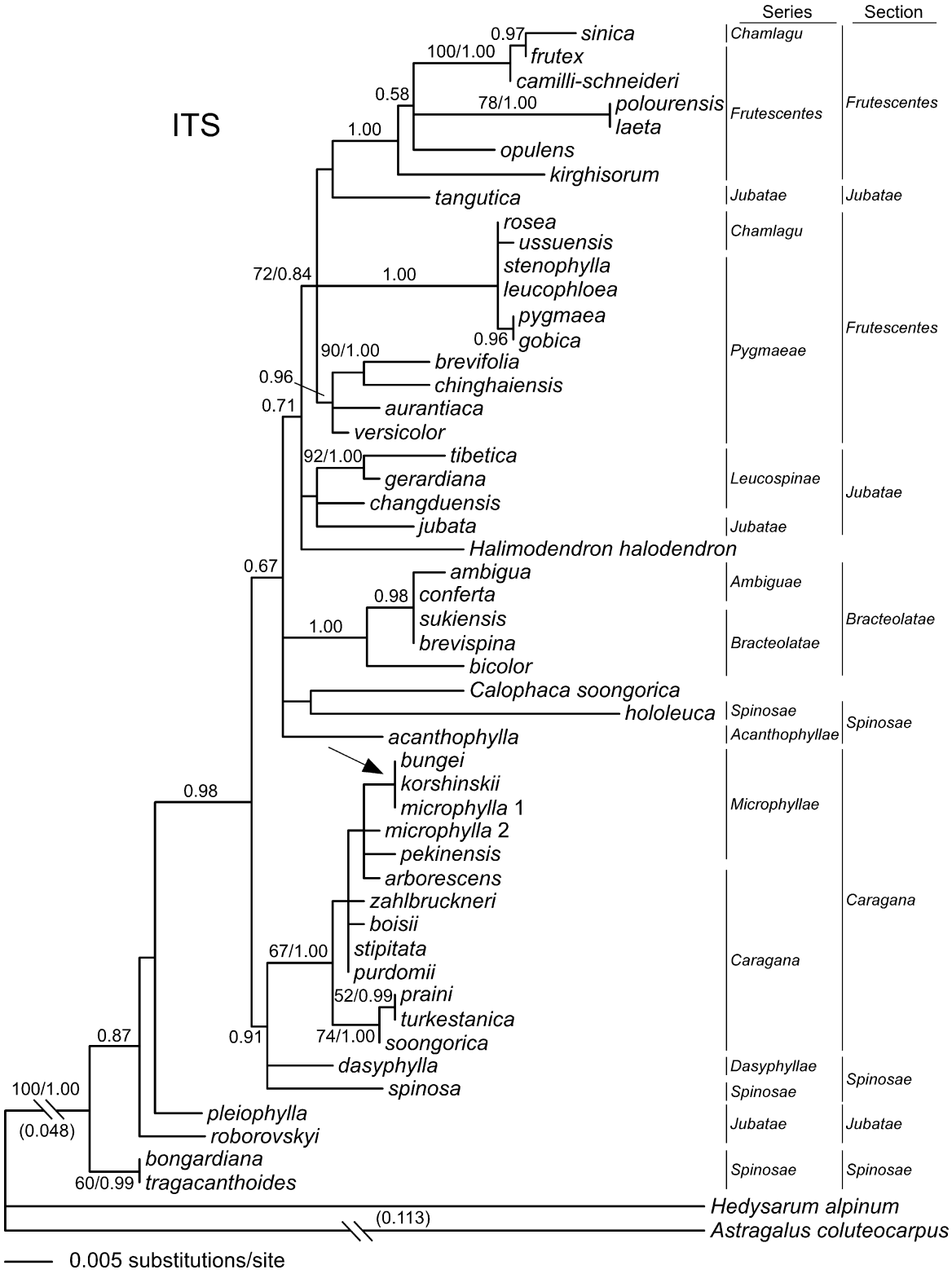

Fig. 1. One of five trees of equal maximum likelihood from phylogenetic analysis of *Caragana* ITS sequence data. Integers are bootstrap values > 50% from a maximum parsimony analysis; decimals are posterior clade probabilities > 0.5 from a Bayesian inference analysis. Arrow indicates the branch that collapses in the strict consensus of the five trees. The values in parentheses are the length values of two branches that are too long to depict accurately in the figure. The classification follows Zhang (1997).

ana + *C. pleiophylla* + *C. roborovskyi* + *C. tragacanthoides* (bt < 50; pP < 0.50). Only two of the series that are represented by more than one terminal are unequivocally monophyletic: ser. *Leucospinae* (bt = 91; pP = 1.00) and ser. *Bracteolatae* (bt < 50; pP = 0.67).

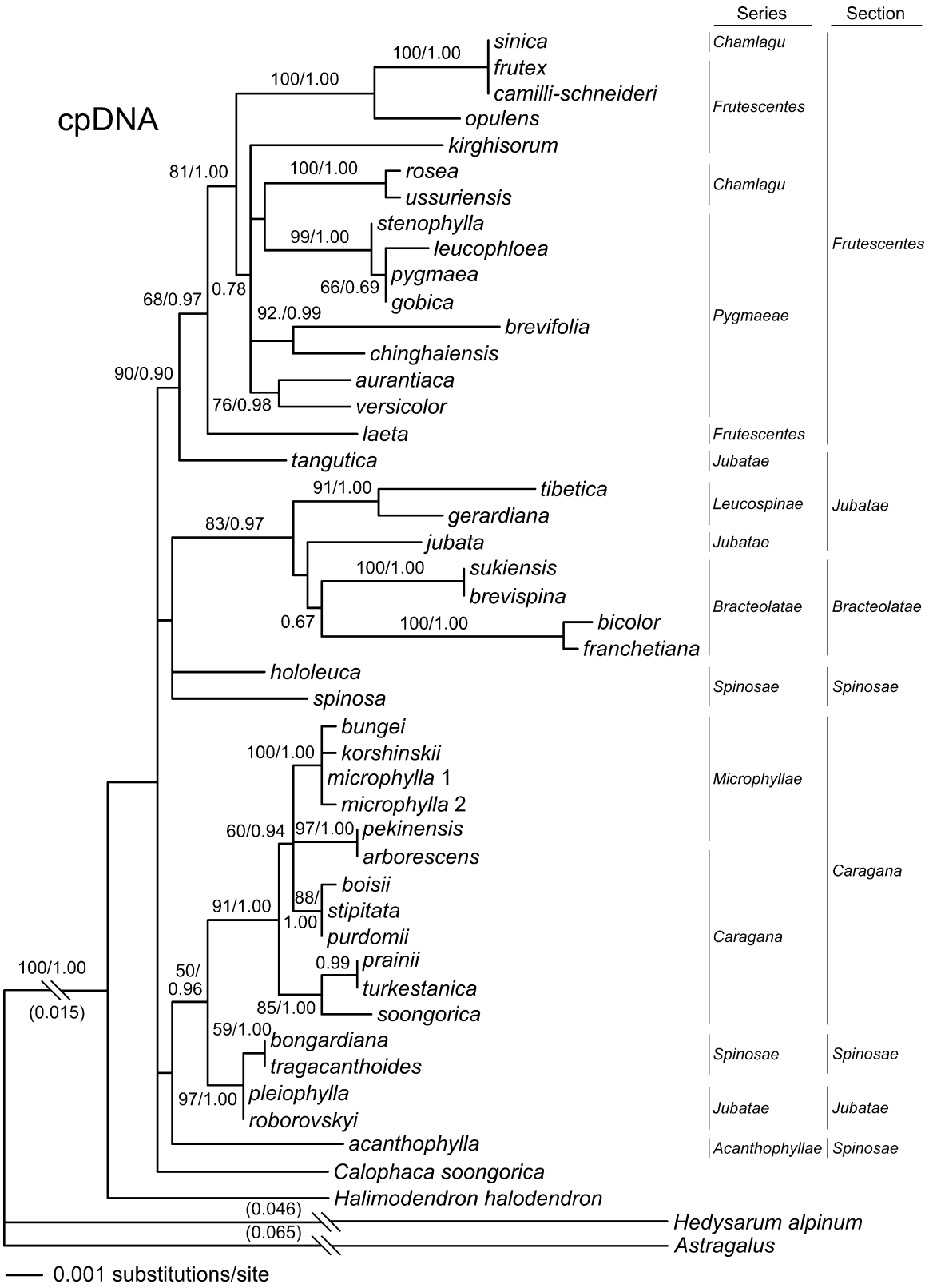

Fig. 2. Maximum likelihood tree of *Caragana* from phylogenetic analysis of combined *trnS–trnG* and *rbcL* (cpDNA) sequence data. Integers are bootstrap values > 50% from a maximum parsimony analysis; decimals are posterior clade probabilities > 0.5 from a Bayesian inference analysis. Values in parentheses are the length values of three branches that are too long to depict accurately in the figure. The classification follows Zhang (1997).

The BI tree differs from the ML tree only in grouping *Calophaca* in an unresolved clade with *Caragana acanthophylla* and *C. hololeuca*, and *Halimodendron* as sister to sect. *Bracteolatae* + sect. *Jubatae* in part (not shown). The only incongruence between the MP tree and the ML tree is that in the MP tree *C. rosea* + *C. ussuriensis* is sister to *C. camilli–schneideri* + *C. frutex* + *C. opulens* + *C. sinica* (not shown).

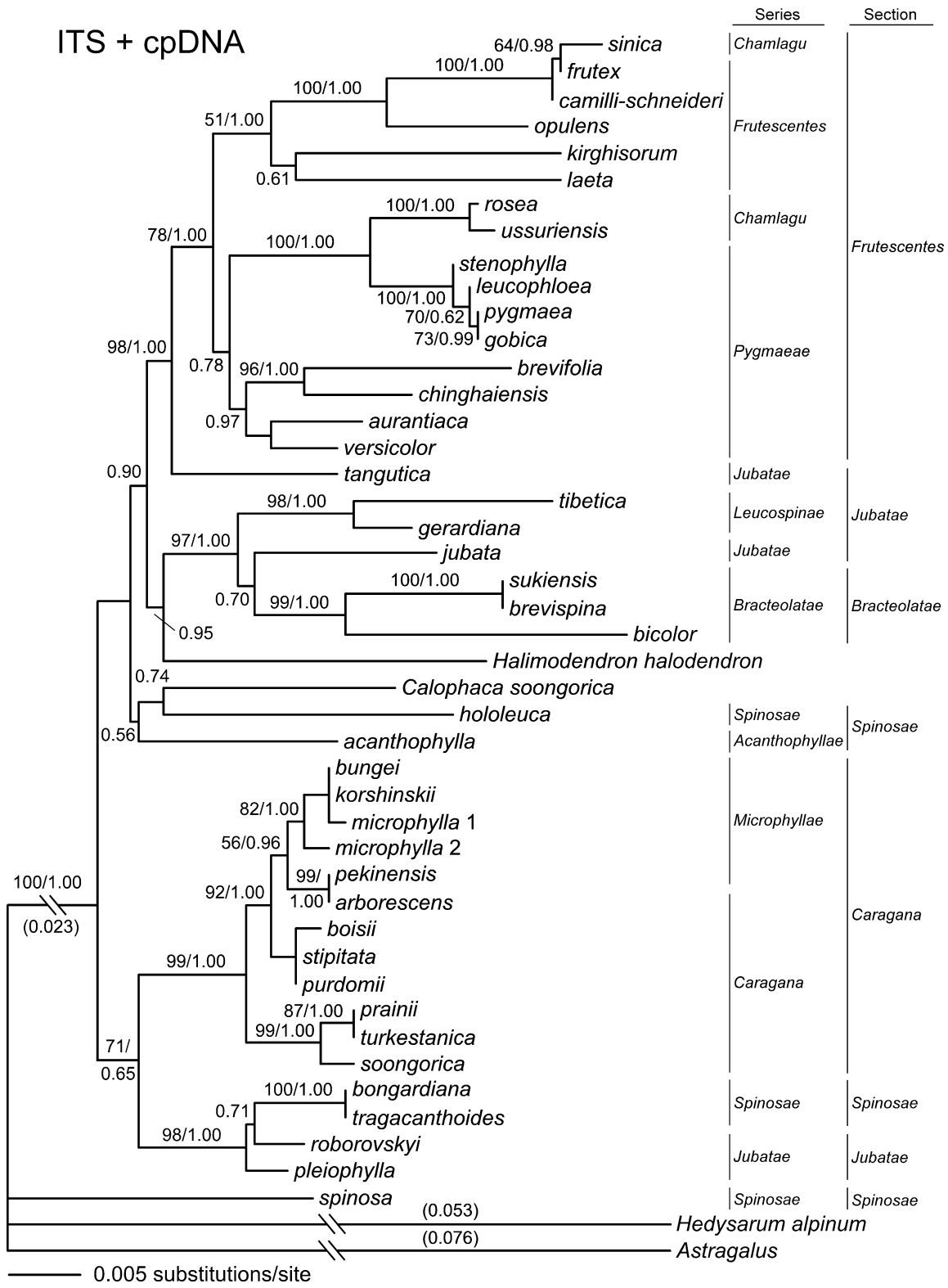

Fig. 3. Maximum likelihood tree of *Caragana* from phylogenetic analysis of combined ITS and cpDNA sequence data. Integers are bootstrap values > 50% from a maximum parsimony analysis; decimals are posterior clade probabilities > 0.5 from a Bayesian inference analysis. Values in parentheses are the length values of three branches that are too long to depict accurately in the figure. The classification follows Zhang (1997).

3.3. Combined 3-gene analysis

The ML analysis resulted in one optimal tree (score = 10005.835). *Calophaca* forms a clade with *Caragana hololeuca* of sect. *Spinosae* (bt < 50; pP = 0.74) that is sister to *C. acanthophylla* of sect. *Spinosae* (bt < 50; pP = 0.56; Fig. 3), and *Halimodendron* groups with a clade comprising sects. *Bracteolatae* and *Jubatae* in part (bt < 50; pP = 0.95). Sections *Caragana*, *Bracteolatae*, and *Frutescentes* are all monophyletic (bt = 99, 99, and 78, respectively; pP = 1.00 for all). The species of sect. *Jubatae* are distributed among three clades: *C. tangutica* is sister to sect. *Frutescentes* (bt = 98; pP = 1.00), *C. jubata* and a clade of *C. tibetica* + *C. gerardiana* group with sect. *Bracteolatae* (bt = 97; pP = 1.00) as successive sister lineages to the latter with weak support, and *C. pleiophylla* and *C. roborovskyi* group with *C. bongariana* and *C. tragacanthoides* of sect. *Spinosae* (bt = 98; pP = 1.00); the latter two species form a clade (bt = 100; pP = 1.00). Only two of the series that are represented by more than one terminal are unequivocally monophyletic: ser. *Leucospinae* (bt = 98; pP = 1.00) and ser. *Bracteolatae* (bt = 99; pP = 1.00).

There are no instances of incongruence between the ML and BI trees (not shown). The MP analysis differs from the ML analysis only in placing *Caragana acanthophylla* as the first-diverging lineage of the ingroup (not shown).

3.4. Morphological evolution over the combined 3-gene molecular tree

The following characters change unambiguously along branches subtending major clades over a representative tree (Fig. 4; all these clades are supported by pP values of 1.00, and most are supported by bt values > 90): sect. *Caragana*: leaf rachis + petiole deciduous (4), leaflet length/width <2 (6), and pedicel articulated at or above middle (8); sect. *Bracteolatae*: leaflet length/width <2 (6) and calyx shape campanulate (9); sect. *Frutescentes*: leaflet arrangement digitate (2), leaflet pair number 2 (3), and leaf rachis + petiole length <1 cm (5); ser. *Microphyllae* + ser. *Caragana* in part: leaflet pair number 5–10 (3); *C. prainii* + *C. soongorica* + *C. turkestanica*: calyx shape campanulate (9); ser. *Chamlagu* in part + ser. *Frutescentes*: pedicel articulated at or above middle (8) and calyx base gibbous (10); *C. rosea* + *C. ussuriensis*: corolla post-anthesis red or amaranth (12) and standard broadly lanceolate or narrowly obovate (13); *C. leucophloea* + *C. pygmaea* + *C. stenophylla*: leaflet length/width > 3 (6); sect. *Bracteolatae* + ser. *Leucospinae* + *C. jubata*: calyx teeth length/tube length ⩾ 1/3 (11); *C. bongardiana* + *C. pleiophylla* + *C. roborovskyi* + *C. tragacanthoides*: leaf rachis of both long and short branches persistent and sclerotic (4) and standard broadly lanceolate or narrowly obovate (13); and *C. bongardiana* + *C. tragacanthoides*: pod inner wall pubescent (18). The various goodness-of-fit character indices demonstrate low overall levels of character consistency and high overall levels of homoplasy in the data (Table 5).

4. Discussion

4.1. Phylogenetic position of Calophaca and Halimodendron

As in previous studies based on far fewer samples (Sanderson and Wojciechowski, 1996; Wojciechowski et al., 2000), our results suggest that *Caragana* is not monophyletic, with both *Calophaca* and *Halimodendron* nested within it. Most statistical support values for the placement of both of the latter genera, however, are low, and in the cpDNA analysis, *Halimodendron* is placed as sister to *Caragana* + *Calophaca* and *Calophaca* is placed in a multichotomy with the clades of *Caragana*. The only strong support for the placement of either *Calophaca* or *Halimodendron* was recovered from the combined 3-gene analysis, with *Halimodendron* as sister to sects.

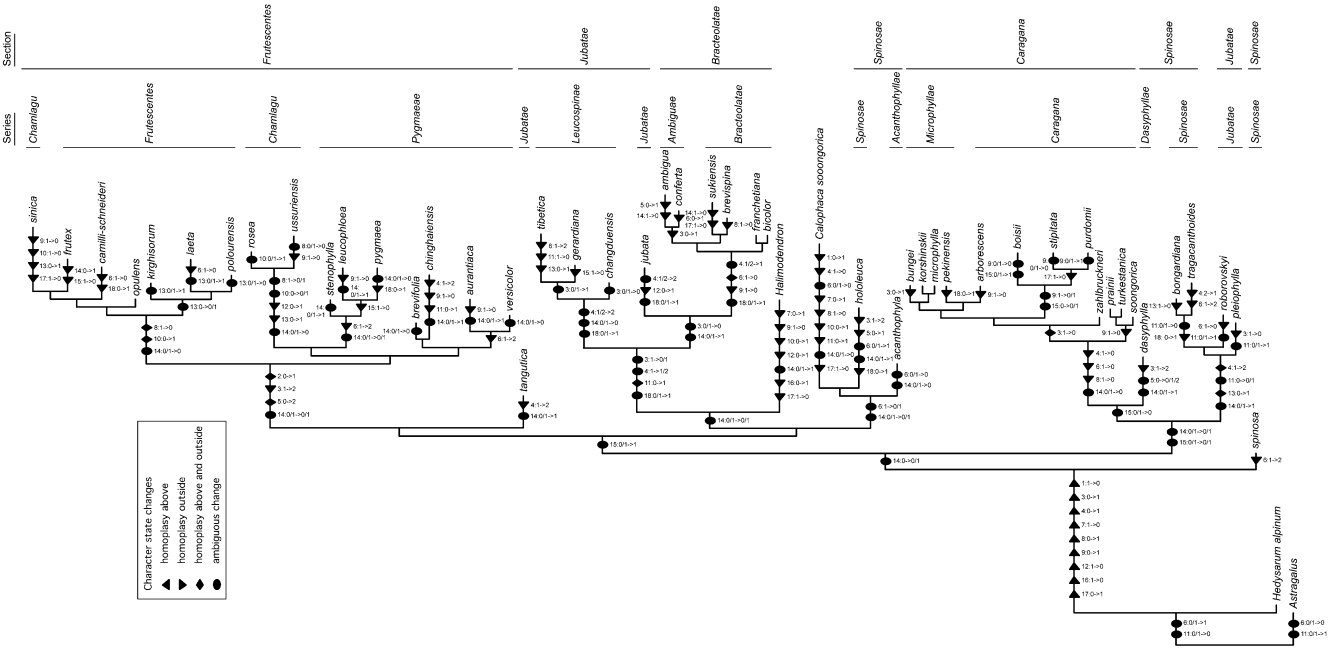

Fig. 4. Eighteen morphological characters used in the infrageneric classification of *Caragana* optimized onto the combined ITS + cpDNA tree with Fitch parsimony. The placement of some taxa is based on either the ITS or the cpDNA results. Black geometric figures indicate characters whose states change unambiguously from one state (number to the left of the arrow) to another (number to right of the arrow). An upward-pointing triangle indicates that the character state also evolves above the branch, a downward-pointing triangle indicates that the character state also evolves below the branch, and a diamond indicates that the character state also evolves both above and below the branch. Circles indicate characters that may have evolved along a particular branch, depending on the character reconstruction; alternative character states are indicated on either side of the slash. No character state changes in this optimization are unique. Polymorphisms within terminals are not shown. The classification follows Zhang (1997).

Table 5
Statistics for 18 morphological characters optimized over the combined-data molecular topology, sorted in order of increasing HI.

Character[a]	Character description	Minimum steps	Tree steps	Maximum steps	CI[b]	RI[c]	HI[d]
1	Leaf type	1	2	3	0.500	0.500	0.500
16	Pod dehiscence	1	2	3	0.500	0.500	0.500
7	Inflorescence type	1	3	4	0.333	0.333	0.667
2	Leaflet arrangement	4	4	17	1.000	1.000	0.750
12	Corolla color post-anthesis	1	4	6	0.250	0.400	0.750
4	Leaf rachis persistence	2	9	24	0.222	0.682	0.778
15	Exine ornamentation	1	5	9	0.200	0.500	0.800
5	Leaf rachis length	9	11	26	0.818	0.882	0.818
10	Calyx base shape	2	6	10	0.333	0.500	0.833
17	Pod neck length	1	6	8	0.167	0.286	0.833
3	Leaflet pair number	6	14	34	0.429	0.714	0.857
11	Calyx teeth length/tube length	1	7	14	0.143	0.538	0.857
13	Standard shape	1	7	9	0.143	0.250	0.857
18	Pod inner wall pubescence	1	7	15	0.143	0.571	0.857
6	Leaflet length/width	3	15	28	0.200	0.520	0.867
8	Articulation position in pedicel	6	11	28	0.545	0.773	0.909
9	Calyx shape	1	12	20	0.083	0.421	0.917
14	Wing auricle length/claw length	1	13	18	0.077	0.294	0.923
Average		2.4	7.7	15.3	0.312	0.592	0.841

[a] Designations are those in Tables 3 and 4.
[b] Consistency index.
[c] Retention index.
[d] Homoplasy index.

Bracteolatae and *Jubatae* in part with a pP value of 0.95; the bt value, however, was <50.

Both *Calophaca* and *Halimodendron* differ from *Caragana* in several basic morphological features. The inflorescences of both genera are racemose, whereas that of *Caragana* is fasciculate, geminate, or solitary-flowered. The leaves of *Calophaca* are imparipinnate, whereas those of *Caragana* and *Halimodendron* are paripinnate, ending in a spine or bristle. Finally, the pod of *Halimodendron* is broadly inflated, versus compressed or linear in *Caragana* and cylindrical or linear in *Calophaca* (Polhill, 1981; Liu et al., in press). The morphological distinctions among the genera are reflected in our estimate of morphological evolution in *Caragana*. In the optimization of morphological changes over the combined 3-gene molecular tree, the branches of both *Calophaca* and *Halimodendron* are long (seven and six unambiguous characters, respectively) relative to all others except for that subtending the clade comprising the three genera (nine) and the branch to *C. sinica* (four; Fig. 4). The consistent morphological distinctions among the genera and low overall molecular support suggest that the nested placement of *Calophaca* and *Halimodendron* within *Caragana* yielded by some of the molecular analyses results from long branch attraction. More molecular data are clearly still needed to help resolve the relationships among these three genera.

4.2. Phylogeny and classification of Caragana

4.2.1. Sections Caragana, Bracteolatae, and Frutescentes

Our analyses strongly support three of the five sections delimited by Zhang (1997): *Caragana*, *Bracteolatae*, and *Frutescentes*. These sections were previously defined in the prior classifications of Zhang (1997) and others on the basis of leaf morphology, with sect. *Caragana* on pinnately arranged leaflets and a deciduous rachis, sect. *Bracteolatae* on pinnately arranged leaflets and a persistent rachis, and sect. *Frutescentes* on digitately arranged leaflets and a persistent rachis. Our inference of morphological character evolution onto the combined 3-gene molecular tree only partly supports these characters as corroborating synapomorphies for these sections. A deciduous rachis of character 4: state 0 (4:0) is a synapomorphy for sect. *Caragana*, and digitately arranged leaflets (2:1) are a synapomorphy for sect. *Frutescentes*; the other unambiguous synapomorphies corroborating these clades involve other characters, i.e., leaflet length/width <2 (6:0) and pedicel articulated at or above the middle (8:0) for sect. *Caragana*, and two pairs of leaflets (3:2) and leaf rachis length <1 cm (5:2) for sect. *Frutescentes*. Section *Bracteolatae*, endemic to the Qinghai–Xizang (Tibet) Plateau and the Himalaya, is defined on neither of these characters because they are both symplesiomorphic for this group. It is instead defined on two other unambiguous characters: a leaflet length/width <2 (6:0), and a campanulate calyx shape (9:0).

Within sect. *Caragana*, ser. *Microphyllae* is monophyletic only upon exclusion of *C. pekinensis* or inclusion of *C. arborescens*, depending on the desired limits of circumscription. Series *Caragana* is paraphyletic, with the clade of *C. boisii*, *C. purdomii*, and *C. stipitata* (northern China to eastern Sichuan) grouping as sister to ser. *Microphyllae* + *C. arborescens*, and the clade of *C. prainii*, *C. soongorica*, and *C. turkestanica* (northern China and eastern Mongolian Plateau) grouping as sister to this larger clade.

None of the series within sect. *Frutescentes* were recovered as monophyletic in our analyses. Pojarkova (1945) segregated series *Chamlagu* from Komarov's (1908) ser. *Frutescentes* by its partly digitate, partly pinnate leaflet arrangement. Komarov (1908) placed *Caragana rosea* in ser. *Frutescentes*, but Zhang (1997) transferred it to ser. *Chamlagu* based on the results of phylogenetic analysis with morphological characters. This series groups in two places in the combined 3-gene analysis, one (*C. sinica*) with the species of ser. *Frutescentes*, the other (*C. rosea* and *C. ussuriensis*) with the species of ser. *Pygmaeae*. Our data support the independent evolution of the polymorphic digitate/pinnate leaf arrangement from the strictly digitate condition in *C. sinica* and *C. ussuriensis*. *Caragana sinica* can be accommodated within ser. *Frutescentes* on the basis of the synapomorphies that unite the clade: pedicel articulated at or above the middle (8:0), and calyx base gibbous (10:1). Although there is strong support for the placement of *C. rosea* and *C. ussuriensis* with the species of ser. *Pygmaeae* on the basis of our molecular data, there is no morphological corroborating evidence for the clade comprising these species. Series *Pygmaeae* appears to have been based only on symplesiomorphic morphological characters.

Patterns of chromosome evolution must be considered tentative because only about 28% of the species have been sampled for chromosome complement. Our data nonetheless suggest that polyploidy in *Caragana* is restricted to the sect. *Frutescentes* clade (*C. frutex*, $2n = 32$, tetraploid; *C. sinica*, $2n = 24$, triploid; *C. stenophylla*, $2n = 32$, tetraploid; *C. ussuriensis*, $2n = 48$, hexaploid) and *C. spinosa* ($2n = 32$). Moore (1968) hypothesized that the triploid and hexaploid species *C. sinensis* and *C. ussuriensis*, respectively, originated through allopolyploid speciation between other unspecified members of ser. *Frutescentes* (including ser. *Pygmaeae* sensu Zhang, 1997) and ser. *Microphyllae*. Separate analyses of our ITS and cpDNA data sets place these species in the clade comprising members of sect. *Frutescentes* only and thus do not support Moore's hypothesis. Rather, the data suggest that these species originated through autopolyploid speciation, possibly with *C. frutex* and *C. rosea*, their respective diploid sister species, as parents.

Although our results strongly support section *Bracteolatae*, there is no support for either of its series, i.e., *Ambiguae* and *Bracteolatae*. In the ITS ML consensus, the two species of ser. *Ambiguae* sampled form a polytomy with two species of ser. *Bractolatae* (pP = 0.98), and this clade is sister to another species of ser. *Bracteolatae* (*Caragana bicolor*). Chloroplast DNA data were not available for any species of ser. *Ambiguae* in our study, so this topology could not be further assessed. The morphological analysis over the molecular tree yielded a synapomorphy for ser. *Ambiguae* (leaflet pair number (2–)3–4; character 3: state 1), but this clade was artificially resolved and thus this character may not be a robust synapomorphy for this group.

Sanchir (1979) and Zhao (1993) proposed that *Caragana arborescens*, with pinnate leaflet arrangement, numerous pairs of leaflets, a deciduous leaf rachis, $2n = 16$, and a distribution in forests of the north-temperate zone, reflects the ancestral stock of the genus. This contrasts with Komarov (1908, 1947), who suggested the same for *C. sinica*, with a polymorphic pinnate/digitate leaflet arrangement, two pairs of leaflets, a leaf rachis of long branches persistent and sclerotic and those of short branches deciduous, and $2n = 24$ (triploid). Our data support neither of these hypotheses, with both species highly nested within the molecular phylogeny with strong statistical support. Moore (1968) considered pinnate leaves with numerous leaflets and deciduous rachises as plesiomorphic within *Caragana*, believing that the species of sect. *Caragana* (as ser. *Caragana*) represented the ancestral type. Our data provide support for pinnate leaves as plesiomorphic within the genus, and suggest that the ancestral states for the other characters are: a leaf rachis of long branches persistent and sclerotic, those of short branches deciduous (4:1), and a leaflet pair number of (2–)3–4 (3:1) rather than numerous (5–10 in our analysis). This combination of character states is rare among the species sampled in our analysis, only occurring in *C. acanthophylla*, *C. ambigua*, *C. conferta*, *C. tragacanthoides*, and *Halimodendron*.

4.2.2. Sections Jubatae and Spinosae

Section (or series) *Jubatae* had been defined narrowly across classifications of the genus, with Komarov (1908) including only *Caragana jubata* and *C. tangutica*, and Pojarkova only these two plus *C. hoplites* Dunn. Later authors, however (e.g., Zhao, 1993; Zhang, 1997), expanded this section to include more species shared among two series, one with glabrous pod inner walls (ser. *Jubatae*), the other with pubescent pod inner walls (ser. *Leucospinae*). Our results consistently yielded a non-monophyletic section *Jubatae* sensu Zhao (1993) and Zhang (1997). Our samples of ser. *Leucospinae* form a strongly supported clade in all our analyses, albeit corroborated only by ambiguous characters (leaf rachis of both long and short branches persistent and sclerotic (4:2), wing auricle length/claw length $\leqslant 1/3$ (14:0), and pod inner wall pubescent (18:1)). *Caragana jubata*, of ser. *Jubatae*, however, is placed with ser. *Leucospinae* plus the species of ser. *Bracteolatae* in a clade with strong support in the combined 3-gene analysis. Our data suggest that combining these taxa into a single section corresponding to this clade is warranted on the corroborating morphological character calyx teeth length/tube length $\geqslant 1/3$ (11:1), and perhaps one or more of the ambiguous characters that might also corroborate the clade (Fig. 4).

As for the other species of ser. *Jubatae* sampled, *Caragana tangutica* is strongly placed as the sister lineage of sect. *Frutescentes* in both the cpDNA and combined 3-gene analyses, and *C. pleiophylla* and *C. roborovskyi* group strongly with a clade comprising two species of sect. *Spinosae* (*C. bongardiana* and *C. tragacanthoides*) in these two analyses. Although the molecular data suggest that reclassifying *C. tangutica* with sect. *Frutescentes* is warranted no corroborative morphological characters for this realignment were detected in our analysis. In contrast, two morphological characters corroborate the molecular placement of *C. pleiophylla* and *C. roborovskyi* (leaf rachis of both long and short branches persistent and sclerotic (4:2), and standard broadly lanceolate or narrowly obovate (13:1); Fig. 4). Other authors of classifications (e.g., Pojarkova, 1945; Liu, 1993) included these species together in ser. *Tragacanthoides* Pojark., but also included other species that fall outside this clade in our results, such as *C. franchetiana*, *C. gerardiana*, *C. hololeuca*, and/or *C. tangutica*.

Other than *Caragana bongardiana* and *C. tragacanthoides*, the phylogenetic placement of the species of sect. *Spinosae* are essentially unresolved in our analyses. *Caragana acanthophylla*, *C. dasyphylla*, *C. hololeuca*, and *C. spinosa* are recovered in basal positions in the three analyses, but their placement often differs among analyses and statistical support is always low. Thus, although it is clear that sect. *Spinosae* is not monophyletic, the extent to which this section will require reclassification is not yet clear.

4.3. Homoplasy in morphological characters

The levels of neither CI nor HI appear to be correlated with particular sets of morphological features (i.e., vegetative, flower, or fruit characters; Table 5). Although overall HI is high in the morphological data, the HI of any particular character does not necessarily reflect its utility for diagnosing major clades. For example, three characters (3, 6, and 11) possess a relatively low CI (0.429, 0.200, and 0.143, respectively) and high HI (0.857, 0.867, and 0.857), but each diagnoses a major clade recovered in the combined 3-gene molecular analysis (sect. *Frutescentes*, sect. *Caragana*, and sect. *Jubatae* in part + sect. *Bracteolatae*) with only one total internal reversal among them. Similarly, the 11-step character 8, with nearly the highest HI among characters (but low CI and RI), diagnoses both sect. *Caragana* and the clade of ser. *Chamlagu* in part + ser. *Frutescentes* without internal reversals.

At least two patterns can explain the utility of highly homoplasious characters for clade diagnosability in *Caragana*. The first is the asymmetric distribution of changes in character states across the tree for some characters. For example, of the nine remaining changes that occur in character 8, five occur within terminals and the rest occur along terminal branches (one ambiguously occurs along a branch subtending two terminals). Thus the high homoplasy value results from a high proportion of shallow changes on the tree. The second is the presence of a rare character state in a multistate character. For example, character 3 has three states, one of which (state 2) diagnoses sect. *Frutescentes* without internal reversals, and otherwise changes only along two terminal branches. The other 11 changes of this 14-step character exclu-

sively involve the other two character states. Although frequency of change has often been assumed to be a critical element in the significance of characters for phylogeny and classification, such changes may not be detrimental and can actually improve the ability to recognize well supported groups (Källerjö et al., 1999; Wenzel and Siddall, 1999; Borsch et al., 2003). This appears to be the case in our study.

4.4. Comparison to a previous study and future phylogenetic work on *Caragana*

The results of the only previous phylogenetic study of *Caragana* based on DNA sequences (Hou et al., 2008) differ from ours in several notable respects. Of the 20 species in the previous study, 14 are included in our analysis. We conducted a MP analysis on a data matrix that included both our ITS sequences and those of the previous study that were downloaded from GenBank and configured to our prior alignment. Of the 14 species in common, nine grouped as sister to our samples of the same species in the strict consensus, whereas five did not (results not shown). The sample of *C. roborovskyi* from the previous study grouped with our and their sample of *C. tibetica* (bt = 99), their *C. bongardiana* grouped with our *C. pekinensis* and their *C. sibirica* Fabr. (bt = 61), their *C. hololeuca* grouped with our *C. bongardiana* and *C. tragacanthoides* (bt = 94), their *C. dasyphylla* was nested within our ser. *Pygmaeae* + *C. rosea* + *C. ussuriensis* clade (bt = 99), and their *C. sinica* grouped as sister to this clade (bt = 88). Other major differences in the results of the previous study compared to ours are as follows. *Caragana bongardiana*, *C. dasyphylla*, and *C. hololeuca* form a clade (bt = 99) that is sister (bt = 99) to the samples of sect. *Frutescentes*. These clades form the sister (bt = 93) of a clade comprising *C. acanthophylla* and *C. bicolor* (bt = 91), and all these clades form the sister (bt = 89) of a clade containing *C. jubata*, *C. roborovskyi*, and *C. tibetica* (bt = 94; Hou et al., 2008: Fig. 1). The differences in tree topology between our ITS results and those of the previous study could be due to one or more of the following: (1) differential sampling, both in genic regions and in species; (2) sequence alignment ambiguities; (3) differences in species identification; and (4) species polyphyly, especially in putative hybrids such as *C. sinica*. An additional complication, however, is that when we included only the ingroup samples of the previous study in an independent analysis of ITS sequences, we recovered a strict consensus that differed substantially from the ITS tree in Hou et al. (2008: Fig. 1). Further assessment of the conflicts between these two studies will require a cross-comparison of sample vouchers and the generation of *rbcL* and *trnL–trnF* sequences for samples not included in one or the other study.

The lack of phylogenetic resolution among the species of sect. *Spinosae* in our results exemplifies the larger problem of non-resolution at the base of the *Caragana* + *Calophaca* + *Halimodendron* clade with our data. This lack of resolution was observed in all three analyses, and so appears not to be a problem inherent to one data set or another, e.g., alignment ambiguities in ITS. More data from other genic regions are clearly required to resolve the base of this clade. Additional studies ideally would include multiple samples of *Calophaca* and *Halimodendron* to assess the monophyly of these two genera, as well as representatives of each of the unresolved lineages based on our analyses.

Because of this basal ambiguity, the development of a revised classification of *Caragana* is premature. The results of this study nonetheless indicate that any reclassification should involve the recognition of sects. *Caragana* and *Frutescentes*, the transfer of the species of ser. *Leucospinae* and *C. jubata* to sect. *Bracteolatae*, and the resurrection of ser. *Tragacanthoides* of Pojarkova (1945) and others (at the sectional level; comprising at least *C. bongardiana*, *C. pleiophylla*, *C. roborovskyi*, and *C. tragacanthoides*). The rest of the genus not readily placed to one of these clades with the morphological synapomorphies indicated in Fig. 4 must remain as incertae cedis until more data can be applied in an attempt to resolve the basal relationships of the genus. Particular focus should be placed on obtaining sequence data from the species of southern Asia that are undersampled for molecular data, such as *C. maimanensis* Rech. f. and *C. ulicina* Stocks from Afghanistan and Pakistan.

Acknowledgments

Thanks to C.Y. Chang for providing *Caragana* field collections from Xinjiang Province, China, Ch. Sanchir and J. Sara for collaboration in the field with the first author in Mongolia, B. Bartholomew for herbarium assistance at the California Academy of Sciences (CAS), M. Wojciechowski for material of *Hedysarum*, and one anonymous reviewer for suggestions in improving the manuscript. Funding from the Chinese Academy of Sciences (KSCX2-YW-2-069, KSCX2-YW-R-136) and China Scholarship Council (CSC) supported M.L. Zhang's six-month visit to the CAS for this study. Most of the laboratory work for this study was conducted in the Center for Comparative Genomics at CAS.

References

Borsch, T., Hilu, K.W., Quandt, D., Wilde, V., Neinhuis, C., Barthlott, W., 2003. Noncoding plastid *trnT–trnF* sequences reveal a well resolved phylogeny of basal angiosperms. J. Evol. Biol. 16, 558–576.

Dieffenbach, C.W., Dveksler, G.S. (Eds.), 1995. PCR Primer: a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Gorbunova, N.N., 1984. De generis *Caragana* Lam. (Fabaceae) notae systematicae. Novosti. Sist. Vyssh. Rast. 21, 92–101.

Hou, X., Liu, J.-E., Zhao, Y.-Z., 2008. Molecular phylogeny of *Caragana* (Fabaceae) in China. J. Syst. Evol. 46, 600–607.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Huelsenbeck, J.P., Rannala, B., 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53, 904–913.

Källerjö, M., Albert, V.A., Farris, J.S., 1999. Homoplasy *increases* phylogenetic structure. Cladistics 15, 91–93.

Komarov, V.L., 1908. Generis *Caragana* monographia. Trudy Glavn. Bot. Sada. 29, 179–388.

Komarov, V.L., 1947. V.L. Komarov Opera Selecta. Academic Science URSS, Moscow and Leningrad. pp. 159–342.

Liu, Y.-X., 1993. *Caragana*. In: Fu, K.T. (Ed.), Flora Reipublicae Popularis Sinicae. Science Press, Beijing, pp. 13–67. 42(1).

Liu, Y.X., Chang, Z.Y., Yakolev, G.P., in press. *Caragana*. In: Wu, Z.Y., Raven, P.H. (Eds.), Flora of China, vol. 10. Science Press, Beijing and Missouri Botanical Garden Press, St. Louis.

Lock, J.M., 2005. Tribe Hedysareae. In: Lewis, G., Schrire, B., MacKinder, B., Lock, M. (Eds.), Legumes of the World. Kew Publishing, pp. 489–495.

Maddison, D.R., Maddison, W.P., 2000. MacClade 4: analysis of Phylogeny and Character Evolution. Version 4.0. Sinauer Associates, Sunderland.

McNeill, J., Barrie, F.R., Burdet, H.M., Demoulin, V., Hawksworth, D.L., Marhold, K., Nicolson, D.H., Prado, J., Silva, P.C., Skog, J.E., Wiersema, J.H., Turland, N.J., 2006. International Code of Botanical Nomenclature (Vienna Code). A.R.G. Gantner, Ruggell, Liechtenstein.

Moore, R.J., 1968. Chromosome numbers and phylogeny in *Caragana* (Leguminosae). Can. J. Bot. 46, 1513–1522.

Olmstead, R.G., Michaels, H.J., Scott, K.M., Palmer, J.D., 1992. Monophyly of the Asteridae and identification of their major lineages inferred from DNA sequences of *rbcL*. Ann. MO Bot. Gard. 79, 249–265.

Pojarkova, A.I., 1945. *Caragana*. In: Komarov, V.L., Schishkin, B.K. (Eds.), Flora of the USSR, vol. 11. Academiae Scientiarum URSS, Moscow and Leningrad, pp. 327–368.

Polhill, R.M., 1981. Galegeae. In: Polhill, R.M., Raven, P.H. (Eds.), Advances in Legume Systematics, Part 1. Royal Botanic Gardens, Kew, pp. 357–363.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sanchir, C., 1979. Genus *Caragana* Lam., systematics, geography, phylogeny and economic significance in study on flora and vegetation of P.R. Mongolia, vol.1. Academic Press, Ulan Bator.

Sanchir, C., 1980. Outline of *Caragana* Lam. species. Institute of Botany, P.R. Mongolia Academy, Ulan Bator, Mongolia (vol. 4, pp. 106–123).

Sanchir, C., 1999. System of the genus *Caragana* Lam. (Fabaceae). Acta Sci. Nat. Univ. Inner Mongol. 30, 501–512.
Sanderson, M.J., Wojciechowski, M.F., 1996. Diversification rates in a temperate legume clade: are there "so many species" of *Astragalus* (Fabaceae)? Am. J. Bot. 83, 1488–1502.
Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., Small, R.L., 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92, 142–166.
Swensen, S.M., Luthi, J.N., Rieseberg, L.H., 1998. Datiscaceae revisited: monophyly and the sequence of breeding system evolution. Syst. Bot. 23, 157–169.
Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland.
Wang, Y.-G., Fritsch, P.W., Shi, S.-H., Almeda, F., Cruz, B.C., Kelly, L.M., 2004. Phylogeny and infrageneric classification of *Symplocos* (Symplocaceae) inferred from DNA sequence data. Am. J. Bot. 91, 1901–1914.
Wenzel, J.W., Siddall, M.E., 1999. Noise. Cladistics 15, 51–64.
Wojciechowski, M.F., 2005. *Astragalus* (Fabaceae): a molecular phylogenetic perspective. Brittonia 57, 382–396.
Wojciechowski, M.F., Sanderson, M.J., Steele, K.P., Liston, A., 2000. Molecular phylogeny of the "temperate herbaceous tribes" of papilionoid legumes: a supertree approach. In: Herendeen, P.S., Bruneau, A. (Eds.), Advances in Legume Systematics, Part 9. Royal Botanic Gardens, Kew, pp. 277–298.
Wojciechowski, M.F., Lavin, M., Sanderson, M.J., 2004. A phylogeny of legumes (Leguminosae) based on analysis of the plastid *matK* gene resolves many well-supported subclades within the family. Am. J. Bot. 91, 1846–1862.
Wu, C.-Y. (Ed.), 1980. Vegetation of China. Science Press, Beijing.
Zhang, M.-L., 1997. A reconstructing phylogeny in *Caragana* (Fabaceae). Acta Bot. Yunnan. 19, 331–341.
Zhang, M.-L., Tian, X.-Y., Ning, J.-C., 1996. Pollen morphology and its taxonomic significance of *Caragana* Fabr. (Fabaceae) from China. Acta Phytotax. Sin. 34, 397–409.
Zhang, M.-L., Huang, Y.-M., Kang, Y., Wang, Y.-W., 2002. Biodiversity and biogeography of legumes in Ordos Plateau. China Bull. Bot. Res. 22, 497–502.
Zhao, Y.-Z., 1993. Taxonomic study of the genus *Caragana* from China. Acta Sci. Nat. Univ. NeiMongol. 24, 631–653.
Zhou, Q.-X., Yang, Y.-P., Zhang, M.-L., 2002. Karyotypes of fourteen species in *Caragana*. Bull. Bot. Res. 22, 492–496.