Published in "Molecular Phylogenetics and Evolution" 62(2): 612-625, 2012
which should be cited to refer to this work.

Allopolyploid origin of highly invasive *Centaurea stoebe* s.l. (Asteraceae)

Patrik Mráz [a,*], Núria Garcia-Jacas [b], Emilie Gex-Fabry [a], Alfonso Susanna [b], Laia Barres [b], Heinz Müller-Schärer [a]

[a] Department of Biology, Unit of Ecology & Evolution, University of Fribourg, Chemin du Musée 10, CH-1700 Fribourg, Switzerland
[b] Botanic Institute of Barcelona (IBB-CSIC-ICUB), Passeig del Migdia, s.n., E-08038 Barcelona, Spain

A B S T R A C T

Spotted knapweed (*Centaurea stoebe*) occurs from Western Asia to Western Europe both as diploid and tetraploid cytotypes, predominantly in single-cytotype populations with higher frequency of diploid populations. Interestingly, only tetraploids have been recorded so far from its introduced range in North America where they became highly invasive.

We performed phylogenetic and network analyses of more than 40 accessions of the *C. stoebe* and *C. paniculata* groups and other related taxa using cloned internal transcribed spacer (ITS) and sequences of the chloroplast *trn*T–*trn*L and *atp*B*rbc*L regions to (i) assess the evolutionary origin of tetraploid *C. stoebe* s.l., and (ii) uncover the phylogeny of the *C. stoebe* group. Both issues have not been studied so far and thus remained controversial.

Cloned ITS sequences showed the presence of two slightly divergent ribotypes occurring in tetraploid cytotype, while only one major ribotype was present in diploid *C. stoebe* s.str. This pattern suggests an allopolyploid origin of tetraploids with contribution of the diploid *C. stoebe* s.str. genome. Although we were not able to detect the second parental taxon, we hypothesize that hybridization might have triggered important changes in morphology and life history traits, which in turn may explain the colonization success of the tetraploid taxon. Bayesian relaxed clock estimations indicate a relatively recent – Pleistocene origin of the tetraploid *C. stoebe* s.l. Furthermore, our analyses showed a deep split between the *C. paniculata* and *C. stoebe* groups, and a young diversification of the taxa within the *C. stoebe* group. In contrast to nrDNA analyses, the observed pattern based on two cpDNA regions was inconclusive with respect to the origin and phylogeny of the studied taxa, most likely due to shared ancient polymorphism and frequent homoplasies.

1. Introduction

Successful biological invasions are generally based on both ecological and evolutionary processes, but the latter ones have been studied to a lesser extent (Vanderhoeven et al., 2010). Ellstrand and Schierenbeck (2000) stressed the importance of hybridization as an evolutionary stimulus of invasiveness. In fact, hybridization between species or between divergent populations within the same species leads to the formations of new genotypic and phenotypic combinations, which may allow the colonization of new ecological niches (Stebbins, 1950; Anderson and Stebbins, 1954). Newly formed plant hybrids are often stabilized by polyploidization (genome doubling) alleviating the problems of chromosome pairing during meiosis and thus sterility (Burke and Arnold, 2001; Abbott et al., 2010). Besides the reproductive assurance of otherwise sterile hybrids, polyploidization leads to fixation of heterotic genotypes and increases genetic variation through a higher number of allelic variants per locus (Comai, 2005). Thus, hybridization and polyploidization either alone, or in concert, may considerably increase the adaptive potential as compared to their ancestors. Many polyploids are successful colonizers of naturally or artificially disturbed habitats (Stebbins, 1985; Ehrendorfer, 1980; Thompson, 1991; Brochmann et al., 2004) and thus polyploidy has been listed in several comprehensive studies focusing on putative determinants of invasiveness in plants. (e.g. Verlaque et al., 2002; Küster et al., 2008; Pyšek et al., 2009). However, these meta-analyses have not distinguished autopolyploids, arising within populations of single species, from allopolyploids, in which interspecific hybridization was involved, most likely because of lack of this information. Nevertheless, knowledge of the evolutionary history of invasive species is crucial for understanding underlying mechanisms of their invasion success.

Several invasive polyploids have recently been found to be of hybridogeneous origin due to progress in molecular biology (Gray et al., 1990; Baumel et al., 2002; Ainouche et al., 2004; Vilatersana et al., 2007; Kim et al., 2008; Jacob and Blattner, 2010). Interspecific hybridization is usually inferred by biparentally inherited

* Corresponding author.
E-mail address: patrik.mraz@unifr.ch (P. Mráz).

nuclear DNA markers, most often using the internal transcribed spacer (ITS). Wide-spread use of ITS1–5.8S–ITS2 stems from their easy amplification with universal primers and relatively high level of interspecific polymorphism (Baldwin et al., 1995). On the other hand, some constraints might hamper its application for phylogenetic reconstructions. More specifically, ITS nrDNA shows higher level of homoplasy than other nuclear markers, the nucleotide position is not independent due to conservative secondary structure, paralogues/ortologues can be frequent and they can be subject to various level of intra-or inter-array homogenization (Álvarez and Wendel, 2003). However, in the case of suppression of the last mentioned mechanism, and thus retention of two (or more) divergent ITS copies within one genome, this marker may prove to be highly informative with respect to the hybridization history. Indeed, many case-studies documented a hybridogeneous origin of polyploid taxa using multicopy ITS or ETS (external transcribed spacer) nrDNA markers (Soltis and Soltis, 1991; Sang et al., 1995; Campbell et al., 1997; Andreasen and Baldwin, 2003; Fehrer et al., 2009). In addition to biparentally inherited polymorphism assessed by nuclear markers, incongruencies between nuclear and plastid phylogenies may indicate reticulation event(s) (Rieseberg and Soltis, 1991; Soltis and Kuzoff, 1995).

Frequent hybridization and polyplodization considerably shaped the evolutionary pattern in the species-rich genus *Centaurea* L. (Hellwig, 2004). Interspecific homoploid hybridization is frequent and single hybrids or hybrid populations are often easily recognized due to their intermediate morphology (e.g. Kummer, 1977; Fernández Casas and Susanna, 1986; Garcia-Jacas and Susanna, 1994; Ochsmann, 2000; Koutecký, 2007; Blair and Huffbauer, 2010; Pisanu et al., 2011). The situation is however more complicated in widespread polyploid taxa. Given their high frequency, surprisingly little is known about their auto- or allopolyploid origin. Based on polysomic segregation at two allozymic loci, an autopolyploid origin has been suggested for the tetraploid cytotype of *C. phrygia* L. (Hardy et al., 2000). On the contrary, a quite complex scenario involving several diploid and tetraploid species in several steps has been proposed for the west Mediterranean tetraploid and hexaploid *C. toletana* Boiss. cytotypes (Garcia-Jacas et al., 2009). Similarly, the presence of different ITS paralogues in *C. boissieri* subsp. *atlantica* (Font Quer) Blanca and *C. debdouensis* Breitw. and Podlech has been explained by recent hybridization (Suárez-Santiago et al., 2007).

The *Centaurea stoebe* group (*Centaurea* sect. *Centaurea*, formerly sect. *Acrolophus*) is represented by diploid ($2n = 2\times = 18$) and tetraploid ($2n = 4\times = 36$) populations occurring sympatrically in the native European range (Ochsmann, 2000; Španiel et al., 2008; Treier et al., 2009). The taxonomic position of both cytotypes is still debated. While Ochsmann (2000) proposed to distinguish diploid (*C. stoebe* subsp. *stoebe*; *C. stoebe* s.str. thereafter) from tetraploid cytotype at the subspecies level [*C. stoebe* subsp. *micranthos* (Gugler) Hayek; *C. stoebe* s.l. thereafter], Španiel et al. (2008) suggested a single species concept with no recognition of intraspecific units. In addition to diploid and tetraploid *C. stoebe* L. several other morphologically similar species to *C. stoebe* have been described [e.g. *C. corymbosa* Pourr., *C. reichenbachii* DC., *C. triniifolia* Heuff., *C. vallesiaca* (DC.) Jord.], but their phylogenetic relationship to *C. stoebe* remains obscure (Ochsmann, 2000).

Centaurea stoebe has been introduced to North America at the end of 20th century and since that became highly invasive (Sheley et al., 1998). More importantly, only the tetraploid cytotype has been recorded so far in the introduced range (Treier et al., 2009; Mráz et al., 2011). This pronounced shift in cytotype composition between the native and introduced range could either be the result of the stochastic introduction of only the tetraploid cytotype, or tetraploids might have a demographic advantage over the diploids, in the case if the diploids had been also introduced (Treier et al., 2009). Both cytotypes differ in their life cycle and this trait could tentatively explain the invasion success of the polyploid cytotype. In fact, tetraploids are short-lived perennials and polycarpic, while diploids are predominantly annual or biennial monocarpic plants (Boggs and Story, 1987; Müller, 1989; Ochsmann, 2000; Story et al., 2001; Henery et al., 2010; Mráz et al., 2011). In addition to the different life cycle, recent multivariate morphometric study based on plants grown under uniform conditions from more than 60 populations from both the native and introduced range showed that the cytotypes also differ in other morphological traits, thus supporting the distinct taxonomic status of both cytotypes (Mráz et al., 2011). Differences in phenotypic and life-cycle traits between cytotypes could be the results of direct polyploidization (autopolyploidization), as whole genome duplication might induce morphological and physiological changes (Müntzing, 1936; Blakeslee, 1941; Maherali et al., 2009), or alternatively, hybridization associated with polyploidization (allo-polyploidization). In the latter case, greater phenotypic and genetic differences between diploid progenitors and their polyploid derivates might be expected than under autopolyploidy, although the extent of differentiation depends on the divergence between the parental taxa (Stelkens et al., 2009).

The aim of the present study was (i) to determine the origin (auto- vs. allopolyploid) of the tetraploid cytotype of *C. stoebe* using biparentally inherited nrDNA ITS marker and two cpDNA loci and (ii) to infer their relationship with closely related taxa belonging to the *C. stoebe* group.

2. Material and methods

2.1. Plant material and ploidy level determination

Forty-two accessions of *Centaurea stoebe* s.l. originating from 38 populations sampled across the native European and introduced North American range and representing all known cytotypes (i.e. $2\times$ and $4\times$ as major cytotypes, and $3\times$ and $6\times$ as rare ones; cf. Mráz et al., 2011) were included in the present study (Table 1). Within *C. stoebe* s.l. we included also the accessions recognized by Ochsmann as separate taxa (e.g. *C. reichenbachii*, *C. tauscherii* A. Kern., *C. triniifolia*, *C. vallesiaca*) to cover variation as large as possible within the group (Ochsmann, 2000). In addition, nine other species were added: *C. cuneifolia* Sm., a species morphologically similar to *C. stoebe* from the Balkans; three species from the *C. paniculata* group, namely *C. aplolepa* Moretti, *C. leucophaea* Jord., and *C. paniculata* L.; and three species showing similar ITS sequences based on a previous study (*C. donetzica* Klokov, *C. sarandinakiae* N.B. Illar., and *C. vankovii* Klokov; see Garcia-Jacas et al., 2006). All taxa belong to the sect. *Centaurea*, except of *C. donetzica*, *C. sarandinakae* and *C. vankovii* which are members of sect. *Phalolepis* (Cass.) DC. (Wagenitz and Hellwig, 1996). Leaf material for DNA extraction was collected either from the plants in the field or from seed-derived plants cultivated in the greenhouse, dried in silica-gel and stored at room temperature. Rarely, herbarium specimens were used. Details on vouchers, population codes and GenBank accession numbers are given in Table 1.

Ploidy level estimations and chromosome counts were determined on seed-derived plants cultivated in the greenhouse, or in rare cases on silica-gel dried material. Details for sample preparation and analyses using flow cytometry and chromosome counting are given in Mráz et al. (2011). Most of the ploidy estimations presented here are new (see Table 1), although some are from our previous publications (Treier et al., 2009; Mráz et al., 2011). For some taxa for which we used already published ITS sequences, ploidy level information were taken from Ochsmann (2000).

Table 1
Origin of the plant material used for the present study, total number of ITS clones sequenced per sample/number of clones used for phylogenetic analyses. A dash indicates failure of cloning (ITS) or not analysed (cpDNA).

Taxon and ploidy[a]	Sampling code	Code used for analyses	Country code, locality, altitude, collector(s), date and herbarium	Total #ITS clones/retained for plots	GenBank accession ITS	cpDNA haplotype	GenBank accession cpDNA (*trnL–trnT*/*atpB–rbcL*)
C. aplolepa 2×	IT2-12	A	IT, Capo Berta, 20 m, Müller-Schärer, 1.6.2008 (NHMR)	5/2	JF913981–JF913982	H2	JF960874/JF960915
C. corymbosa 2×*		CO	FR, Narbonne, La Clappe, M. Riba, 1995 (BC).	Not cloned		–	–
C. cuneifolia 2×	BG7-1	CU	BG, Belovo, 443 m, P. Mráz and Mrázová, 3.8.2008 (NHMR)	8/3	JF913983–JF913985	–	–
C. diffusa 2×*	DIF	DI	ARM, Talin, between villages Pokr Arthik and Bagravan, Susanna 1589 et al., 26.8.1995 (BC).	Not cloned	–	–	–
C. donetzica 2×	DON	DO	UA, Donetzk region, Krasny Liman, Romaschenko, 13.07.2009 (BC).	5/3	JF913986–JF913988	–	–
C. leucophea 2×	FRE-4	L	FR, Drôme, Allan, 283 m, Treier and Broenniman, 4.8.2005 (NHMR)	4/1	JF913989	H3	JF960872/JF960913
C. paniculata 2×	FRA15-2	P	FR, Ain, St-Maurice de Gourdans, 190 m, P. Mráz and Priestman, 2.5.2008 (NHMR)	5/2	JF913990–JF91399	H9	JF960869/JF960910
C. sarandinakiae 4×	SARAN	SA	UA, Crimea, Kara-Dag, Futorna and Romaschenko, 13.07.2009 (BC).	4/3	JF913992–JF913994	–	–
C. stoebe 2×	BG1-3	S1	BG, Bosnek, 875 m, P. Mráz and Mrázová, 1.8.2008 (NHMR)	8/2	JF913995–JF913996	H1	JF960856/JF960897
C. stoebe 2×	BG4-4	S2	BG, Dagonovo, 806 m, P. Mráz and Mrázová, 2.8.2008 (NHMR)	7/–	–	H10	JF960859/JF960890
C. stoebe 2×	BG5-5	S3	BG, Yagoruda, Mt. Granchar, 2150 m, Mrázová, 2.8.2008 (NHMR)	8/1	JF913997	H2	JF960860/JF960891
C. stoebe 2×*[b]	SW4-6	S4	CH, Wallis, Ausserberg, 924 m, Thébault, 8.9.2005 (NHMR)	8/1	JF913998	H1	JF960891/JF960932
C. stoebe 2×*	SCHA-10	S5	CH, Graubünden, Ramosch, 1237 m, Treier and Normand, 18.8.2005 (NHMR)	12/2	JF913999–JF914000	H2	JF960880/JF960921
C. stoebe 2×*	DE6-14	S6	DE, Sachsen-Anhalt, Halle, 85 m, Thébault and Broennimann, 21.8.2005 (NHMR)	8/1	JF914001	H4	JF960862/JF960903
C. stoebe 2×*	DE10-3	S7	DE, Sachsen-Anhalt, Zadel, 145 m, Thébault and Broennimann, 22.8.2005 (NHMR)	7/2	JF914002–JF914003	H1	JF960863/JF960904
C. stoebe 2×*[c]	Albida-1	S8	FR, Gard, Anduze, Tisson, 160 m, 5.1995 (NHMR)	8/2	JF914004–JF914005	H8	JF960855/JF960896
C. stoebe 2×[d]	FRA11-3	S9	FR, Haute Loire, Leotoing, 612 m, P. Mráz and Priestman, 1.5.2008 (NHMR)	5/2	JF914006–JF914007	H2	JF960865/JF960906
C. stoebe 2×[d]	FRA13-4	S10	FR, Haute Loire, Espoly-St-Marcel de l'Ermitage, 786 m, P. Mráz and Priestman, 1.5.2008 (NHMR)	5/1	JF914008	H2	JF960867/JF960908
C. stoebe 2×[d]	FRA14-5	S11	FR, Puy de Dôme, Mt. Puy de Crouël, 394 m, P. Mráz and Priestman, 1.5.2008 (NHMR)	6/1	JF914009	H2	JF960868/JF960909
C. stoebe 2×	FRA16-21	S12	FR, Savoie, Termignon, 1398 m, P. Mráz and S. Mráz, 10.9.2008 (NHMR)	7/2	JF914010–JF914011	H3	JF960870/JF960911
C. stoebe 2×	FRA17-1	S13	FR, Savoie, Modane, 1258 m, P. Mráz and S. Mráz, 10.9.2008 (NHMR)	9/2	JF914012–JF914013	H3	JF960871/JF960912
C. stoebe 2×	IT3-16	S14	IT, Piemont, Casellete, 442 m, P. Mráz and S. Mráz, 11.9.2008 (NHMR)	–/–		H5	JF960875/JF960916
C. stoebe 2×[e]	RO14-1	S15	RO, Cheile Turzei, 564 m, P. Mráz and Mrázová, 7.8.2008 (NHMR)	12/3	JF914014–JF914016	H2	JF960877/JF960918
C. stoebe 2×	SER5-2	S16	RS, Vranje, 450 m, P. Mráz and Mrázová, 30.7.2008 (NHMR)	6/3	JF914017–JF914019	H2	JF960883/JF960924
C. stoebe 2×	SER8-3	S17	RS, Mt. Pirot, 1335 m, P. Mráz and Mrázová, 30.7.2008 (NHMR)	14/2	JF914020–JF914021	H1	JF960885/JF960926
C. stoebe 2×	SER9-1	S18	RS, Basara, 952 m, P. Mráz and Mrázová, 30.7.2008 (NHMR)	4/1	JF914022	H12	JF960886/JF960927
C. stoebe 2×*	SRUG-4	S19	RU, Samara, Perevoloki, 78 m, Naumoff, 24.9.2006 (NHMR)	6/2	JF914023–JF914024	H1	JF960887/JF960928
C. stoebe 2×*	SRUG-12	S20	RU, Samara, Perevoloki, 78 m, Naumoff, 24.9.2006 (NHMR)	9/2	JF914025–JF914026	H4	JF960888/JF960929
C. stoebe 2×*	SRUO-2	S21	RU, Dagestan Republic, Karabudokhkentskiy district, 970 m, Nikolaeva, 20.7.2006 (without voucher)	11/3	JF914027–JF914029	H4	JF960889/JF960930
C. stoebe 2×	DK2-421	S22	SK, Devínska Nová Ves, 191 m, P. Mráz and Mrázová, 12.8.2008 (NHMR)	3/1	JF914030	–	–
C. stoebe 2×*	SUAI-2	S23	UA, Poltava, Chutove, 131 m, Treier and Broennimann, 15.9.2005 (NHMR)	8/2	JF914031–JF914032	–	–
C. stoebe 3×	Ma-134	S24	AT, Marchegg, 35 m, Bowman and Farkas, 2007 (NHMR)	5/2	JF914033–JF914034	H1	JF960879/JF960920
C. stoebe 3×	DK2-293	S25	SK, Devínska Nová Ves, 206 m, P. Mráz and Procházka, 15.8.2008 (NHMR)	–/–	–	H6	JF960864/JF960905
C. stoebe 4×	BG1-5	S26	BG, Bosnek, 875 m, P. Mráz and Mrázová, 1.8.2008 (NHMR)	13/3	JF914035–JF914037	H2	JF960857/JF960898
C. stoebe 4×	BG2-2	S27	BG, Topolnitsa, 650 m, P. Mráz and Mrázová, 1.8.2008 (NHMR)	6/2	JF914038–JF914039	H7	JF960858/JF960899
C. stoebe 4×	BG6-1	S28	BG, Yundola saddle, 1638 m, P. Mráz and Mrázová,	7/3	JF914040–	H1	JF960861/JF960902

Table 1 (*continued*)

Taxon and ploidy[a]	Sampling code	Code used for analyses	Country code, locality, altitude, collector(s), date and herbarium	Total #ITS clones/retained for plots	GenBank accession ITS	cpDNA haplotype	GenBank accession cpDNA (*trnL–trnT*/*atpB–rbcL*)
			2.8.2008 (NHMR)		JF914042		
C. stoebe 4×	FRA12-1	S29	FR, Allier, Moulins, 207 m, P. Mráz and Priestman, 31.4.2008 (NHMR)	7/2	JF914043–JF914044	H1	JF960866/JF960907
C. stoebe 4×[f]	HU11-8	S30	HU, Tököl, 107 m, P. Mráz and Mrázová, 27.7.2008 (NHMR)	12/2	JF914045–JF914046	H1	JF960873/JF960914
C. stoebe 4×[g]	RO11-5	S31	RO, Băile Herculeane, 229 m, P. Mráz and Mrázová, 5.8.2008 (NHMR)	13/2	JF914047–JF914048	H1	JF960876/JF960917
C. stoebe 4×	RO14-8	S32	RO, Cheile Turzei, 564 m, P. Mráz and Mrázová, 7.8.2008 (NHMR)	5/3	JF914049–JF914051	H1	JF960878/JF960919
C. stoebe 4×	SER2-3	S33	RS, Bogutovac, 245 m, P. Mráz and Mrázová, 29.7.2008 (NHMR)	7/2	JF914052–JF914053	H2	JF960881/JF960922
C. stoebe 4×	SER4a	S34	RS, Massif of Kopaonik, 1460 m, P. Mráz and Mrázová, 29.7.2008	–/–	–	H1	JF960882/JF960923
C. stoebe 4×	SER7-5	S35	RS, Grkinja, 279 m, P. Mráz and Mrázová, 30.7.2008 (NHMR)	13/3	JF914054–JF914056	H1	JF960884/JF960925
C. stoebe 4×*	SUAD-10	S36	UA, Chernivtsi, 313 m, Treier and Broennimann, 9.9.2005 (NHMR)	6/3	JF914057–JF914059	H1	JF960890/JF960931
C. stoebe 4×	DK-347	S37	SK, Devínska Nová Ves, 232 m, P. Mráz and Procházka, 12.8.2008 (NHMR)	7/2	JF914060–JF914061	–	–
C. stoebe 4×*	USMT10-4	S39	US, Montana, Missoula, 1146 m, Treier and Broennimann, 13.10.2005 (NHMR)	14/2	JF914062–JF914063	H1	JF960893/JF960934
C. stoebe 4×*	USOR10-1	S40	US, Oregon, Klamath Falls, 1263 m, Treier and Broennimann, 13.10.2005 (NHMR)	6/3	JF914064–JF914066	H2	JF960894/JF960935
C. stoebe 4×*	USWI1-3	S41	US, Wisconsin, Necedah, 277 m, Hufbauer, 13.10.2005 (NHMR)	6/3	JF914067–JF914069	H11	JF960895/JF960936
C. stoebe 6×*	URS3	S42	CA, British Columbia, Elko, 920 m, Bourchier, 17.10.2007 (NHMR)	7/3	JF914070–JF914072	H1	JF960892/JF960933
C. vankovii 2×	VAN	V	UA, Crimea, Mt. Demerdji, 1200 m, Futorna and Romaschenko, 15.07.2009 (BC).	4/3	JF914073–JF914074	–	–

ARM – Armenia, AT – Austria, BG – Bulgaria, CA – Canada, CH – Switzerland, DE – Germany, FR – France, HU – Hungary, IT – Italy, RO – Romania, RS – Serbia, RU – Russia, SK – Slovakia, UA – Ukraine, US – United States.

[a] Ploidy level estimations of the plants/populations marked by asterisk (*) are based on previously published data (see Section 2), those without asterisk are new records.

[b] Morphologically corresponds to *C. vallesiaca* DC.

[c] Morphologically corresponds to *C. maculosa* subsp. *albida* (Lecoq and Lamotte) Dostál, det. J.M. Tison.

[d] Morphologically corresponds to *C. maculosa* Lam.

[e] From this site both *C. stoebe* s.l. and *C. reichenbachii* has been reported (cf. Ochsmann, 2000).

[f] Locus classicus of *C. tauscheri* A. Kern.

[g] Locus classicus of *C. triniifolia* Heuf.

2.2. DNA extraction, amplification, cloning, and sequencing

Total DNA was extracted from 10 to 15 mg of silica-dried leaf tissue with the DNeasy 96 Plant Kit (Qiagen Inc., Valencia, CA, USA), or in some cases using CTAB method (Doyle and Doyle, 1987). The ITS region was amplified using the primers 17SE and 26SE (Sun et al., 1994) in a 25 μl reaction volume containing 3 μl of diluted genomic DNA, 10× AmpliTaq buffer, 2.5 mM MgCl2, 0.2 mM dNTPs, 0.2 μM of each primer, 0.5 μl of DMSO (Sigma–Aldrich, St. Luis, MO, USA) and 0.5 U of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, CA, USA). The cycling profile included an initial denaturation step at 94 °C/2 min followed by 35 cycles of 94 °C/1 min 30 s, 57 °C/2 min, 72 °C/3 min, and ended with 72 °C/15 min and 4 °C thereafter.

The *trnT–trnL* locus was amplified using "a" and "b" primers of Taberlet et al. (1991) and the *atpB–rbcL* locus with the primers proposed by Chiang et al. (1998). The PCRs were performed in 25 μl volume containing 5 μl of genomic DNA (4 ng μl^{-1}), 10× PCR Buffer, 1 mM MgCl2, 0.2 mM dNTPs, 0.4 μM of each primer, 0.25 μM BSA and 1 U of Taq polymerase (Qiagen). The cycle profile included the initial denaturation at 94 °C/3 min followed by 36 cycles of 94 °C/30 s, 50 °C/30 s, 72 °C/1 min, and ended with 72 °C/5 min and 4 °C thereafter.

The cloning of ITS1–ITS2 regions was performed using TOPO TA Cloning ® Kit for Sequencing (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol. When possible, at least eight positive colonies from each reaction were screened with direct PCR using T7 and M13 universal primers and following reaction conditions: initial denaturation at 94 °C/10 min followed by 30 cycles of 94 °C/30 s, 55 °C/1 min, and ended with 72 °C/10 min and 4 °C thereafter. Five to fifteen clones per accession were selected for sequencing. Direct sequencing was performed using BigDye Terminator Cycle Sequencing v3.1 (Applied Biosystems, Foster City, CA, USA) following the manufacturer's protocol on an ABI 3730xl capillary sequencer (Applied Biosystems) at the University of Florida ICBR Core Facility. Sequences were edited manually using BioEdit 7.0.5.3 (Hall, 1999) and assembled using Mega 4.01 (Tamura et al., 2007).

2.3. Phylogenetic and network analyses

We used Bayesian and parsimony analyses to infer ITS phylogeny and a distance network analysis (split graphs) from cloned ITS sequences, to which we added previously published sequences of two closely related taxa, *C. corymbosa* and *C. diffusa* Lam. (Garcia-Jacas et al., 2006). The same clones of the same accession, and the clones showing unique substitutions within and between accessions were excluded from the analyses, as they may represent random PCR errors (Cline et al., 1996; Popp and Oxelman, 2001). In some cases, we regrouped slightly different clones obtained from the same accession (differing usually by one substitution) to obtain one consensual sequence and thus to reduce the number of clones for analyses. The data matrix is available on request from the corresponding author. Bayesian posterior probabilities were estimated using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). The evolutionary model was selected using jModeltest (Posada, 2008) and maximum likelihood parameters were specified

according to the Akaike Information Criterion (AIC: TM3ef + g). Bayesian analysis was initiated with random starting trees and continued until the value of standard deviation of split sequences droped below the 0.01 as convergence diagnostic value. Log-output file was monitored with Tracer 4.1 to ensure that all parameters achieved sufficient sampling size (>200). The fraction of the sampled values discarded as burn-in was set at 0.25. Posterior probabilities (PP) of 0.95–1.00 were considered statistically significant. For the Neighbor-Net analyses we used the Neighbor-Net (NN) algorithm (Bryant and Moulton, 2004) as implemented in SplitsTree 4.10 software (Huson and Bryant, 2006) with the criterion set to uncorrected pair-wise (p) distances, excluding both constant and non-informative characters. The haplotype network based on substitution polymorphisms in two assembled cpDNA loci (*trnT–trnL* and *atpB–rbcL*) was constructed using the median-joining algorithm implemented in Network 4.6.0.0 (www.fluxus-engineering.com, Bandelt et al., 1999). The same weight (10) was attributed to all variables sites.

2.4. Divergence time estimation

To estimate the divergence time between different ITS clades, we applied two methods. (i) A molecular clock approach using two slightly different substitution rates. The first one corresponds to 2.51×10^{-9} substitution per site and year based on independently calibrated ITS data of herbaceous representatives of the genus *Eupatorium* (Asteraceae, subtribe Eupatorieae; cf. Schmidt and Schilling, 2000; Kay et al., 2006). The second one corresponds to 3×10^{-9} substitution derived from the tribe Madieae (Baldwin and Sanderson, 1998) and used by Suárez-Santiago et al. (2007) to estimate divergence time between sections *Willkommia* and *Centaurea* (syn. *Acrolophus*). Although the likelihood ratio test (baseml package in PAML; Yang, 2007) rejected the assumption of constancy rate of ITS evolution in our data ($p = 0.01$), this was not the case for data set of Suárez-Santiago et al. (2007) probably due to larger number of taxa used in latter study. The Tamura-Nei substitution model with gamma distribution was chosen using online FindModel server (www.hiv.lanl.gov/content/sequence/findmodel/findmodel) and was applied to calculate mean genetic distances between the major clades using MEGA 4.01. (ii) Dating analyses were also performed with BEAST 1.6.1 (Drummond and Rambaut, 2007) using a Bayesian method based on a relaxed molecular clock hypotheses, implying that the evolutionary rate is not constant over time. Because of the lack of *C. stoebe* fossils, we used two age estimations for a Cardueae tribe phylogeny calibrated with five fossils (Barres et al., unpubl. data) as calibration points. *Centaurea lingulata* Lag. from *Centaurea* subgen. *Cyanus* and *Centaurea behen* L. from subgen. *Centaurea* were added as external calibration points to the *C. stoebe* dataset. *Rhaponticoides hajastana* (Tzvel.) M. V. Agab. & Greuter and *Psephellus persicus* (DC.) Wagenitz from the Centaureinae were also included as most external groups and coded as outgroups. The split of subgen. *Cyanus* from all other Cenatureinae was estimated at c. 13.06 mya (Barres et al., unpubl. data) and was used to calibrate the split of *C. lingulata*. The split of subgen. *Centaurea* was estimated in 10.16 mya (Barres et al., unpublished data) and was used to calibrate the split of *C. behen* from the *C. paniculata–C. stoebe* clade. We used a normal distribution prior with ±1 SD for both calibration points, as they were published age estimations. For the datation dataset, model selection was performed with MrModelTest 2.3 (Nylander, 2004) following the Akaike criteria. The best-fit model selected for the estimation of divergence times was SYM + G. We assumed a Constant Size Coalescent Model for the tree prior and an uncorrelated lognormal distribution for the molecular clock model (Drummond et al., 2006; Ho, 2007). For all other parameters we used the default prior distributions. We ran MCMC chains for 50 million generations. The 10% of the first sampled trees were removed as burn-in, and the posterior probability density was summarized using TreeAnnotator 1.6.1 (Drummond and Rambaut, 2007). Parameter estimates and their 95% highest posterior density intervals (HPDs) are shown in Table 3.

3. Results

3.1. Phylogenetic and Neighbor-Net analyses of the cloned ITS sequences

In total, we sequenced 342 clones, from which 96 clones representing 45 accessions were retained for further analyses. The cloning failed in three accessions (see Table 1). Furthermore, one accession (S2; BG4-4) was excluded from the analyses as we obtained completely different sequences containing many autoapomorphic changes not shared with any other population of *C. stoebe* or other *Centaurea* species. Total alignment of ITS1–5.8 rDNA–ITS2 was 634 bp and ranged from 632 to 634 bp per accession. The number of different ITS clones found in an individual plant varied from 1 to 3. Bayesian phylogenetic analysis revealed obvious separation of the clones of three accessions belonging to the *C. paniculata* group from the remaining clones/sequences (Fig. 1). Within the *C. paniculata* branch, we found two closely related ITS copies consistently differentiated by two substitutions (Figs. 1 and 2). While two sequenced species of *C. paniculata* and *C. aplolepa* showed intra-individual polymorphism by sharing these two ribotypes, only one paniculata ribotype was found in *C. leucophaea*. The clones of the *C. stoebe* group, as well as of other taxa formed one well supported, but highly polytomic clade. Within this major clade another one supported subclade emerged. The major clade and subclade were representing by two divergent ITS sequences here referred to ribotypes (copies) A and B (Fig. 1). The clones belonging to A and B ribotypes differed, with some exceptions, consistently in three substitutions at positions 77, 199 and 577 (Table 2). Three clones (S19b, S21a, S23b) from three diploid *C. stoebe* s.str. accessions, which exhibited A ribotype had at position 199 C instead of T. Furthermore, one clone of tetraploid *C. sarandinakiae* (SAa) belonging to the ribotype B had G and T at positions 77 and 199 respectively, like the ribotype A. Finally, the clone S35c from tetraploid *C. stoebe* s.l. belonging to the B ribotype had at position 77 G instead of A (Table 2). In addition to three diagnostic sites, approximately half of the clones of the B ribotype showed further substitution (C → T) at position 499. This substitution occurred also in one S5a clone (diploid *C. stoebe* s.str. from Switzerland) belonging to the ribotype A. Ribotype A was found in all sequenced diploid, triploid, tetraploid and hexaploid accessions of *C. stoebe* s.l., and in all other diploid taxa, except one tetraploid accession of *C. stoebe* s.l. from Serbia (S35), where only B ribotype and one putatively recombinant clone was found (Fig. 2). Some further subclades showed fairly high support within clade A. Three clones of three diploid *C. stoebe* s.str. accessions (S19, S21, S23) originating from the same region (Russia and Ukraine) were clustered together, and further three clones of accessions S5, S16 and S17 formed another well-supported subclade, however with no clear geographic affinities (two accessions from Serbia, one from Switzerland). Within the A clade the clones of diploid *C. donetzica* and *C. vankovii* formed a further subclade, but with lower support. In contrast to ribotype A, the B ribotype was found in all sequenced polyploid accessions of *C. stoebe* s.l. (i.e. tri-, tetra- and hexa-) and tetraploid C. *sarandinikiae* but was absent in all diploid *C. stoebe* s.str.

The topology of the parsimony strict consensus tree (not shown; RI = 0.8537, CI = 0.6104) was coincident with the Bayesian tree, but bootstrap support values were very low. The

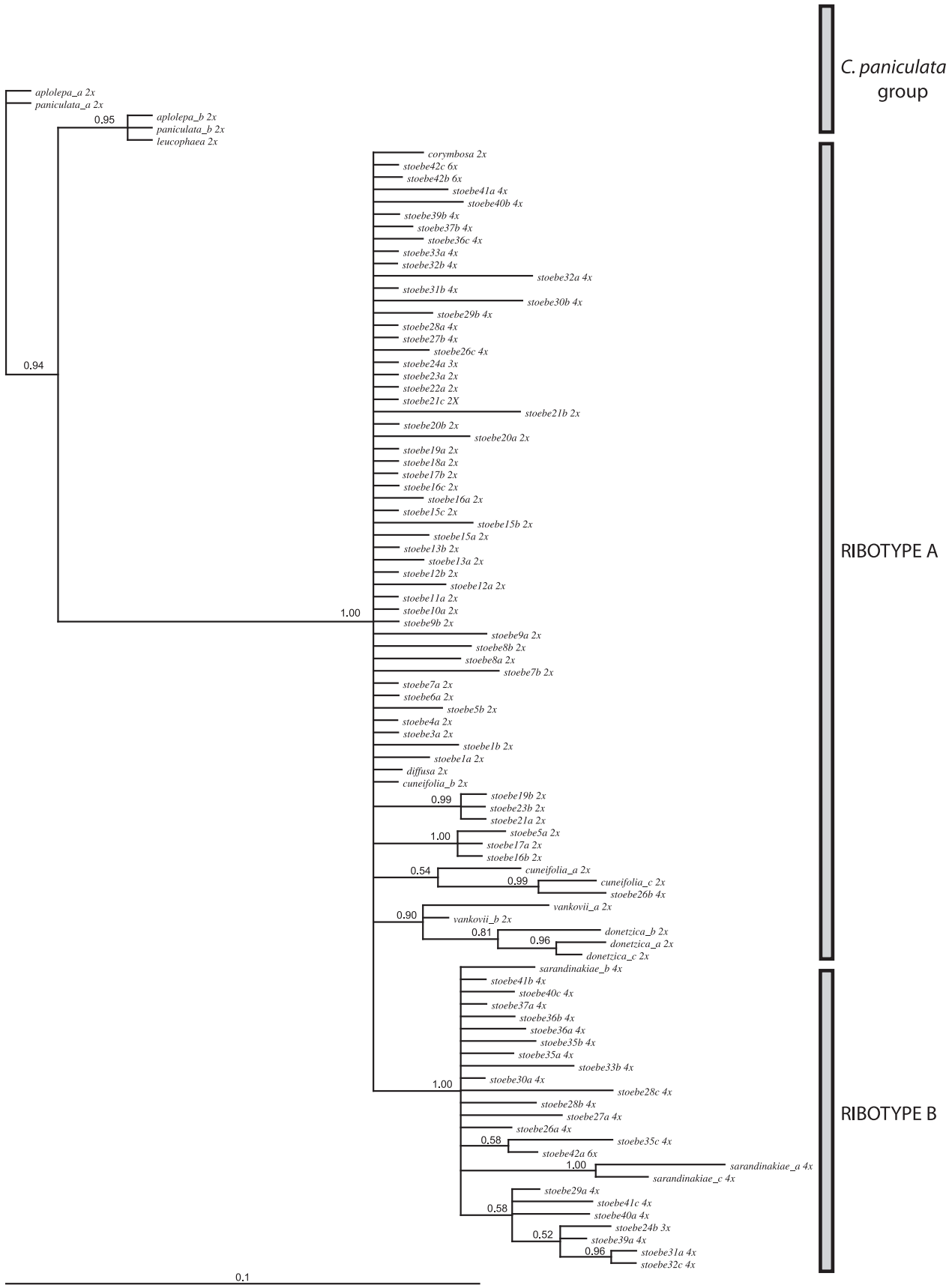

Fig. 1. Bayesian majority rule consensus tree from the ITS data using *Centaurea aplolepa*, *C. leucophea* and *C. paniculata* as outgroup species. Numbers above branches indicate Bayesian-credibility values (PP). Plant codes are as follow: A – *C. aplolepa*; CO – *C. corymbosa*; CU – *C. cuneifolia*; DI – *C. diffusa*; DO – *C. donetzica*; L – *C. leucophea*; P – *C. paniculata*; SA – *C. sarandinakiae*; S1 to S42 – *C. stoebe*; V – *C. vankovii* (see Table 1).

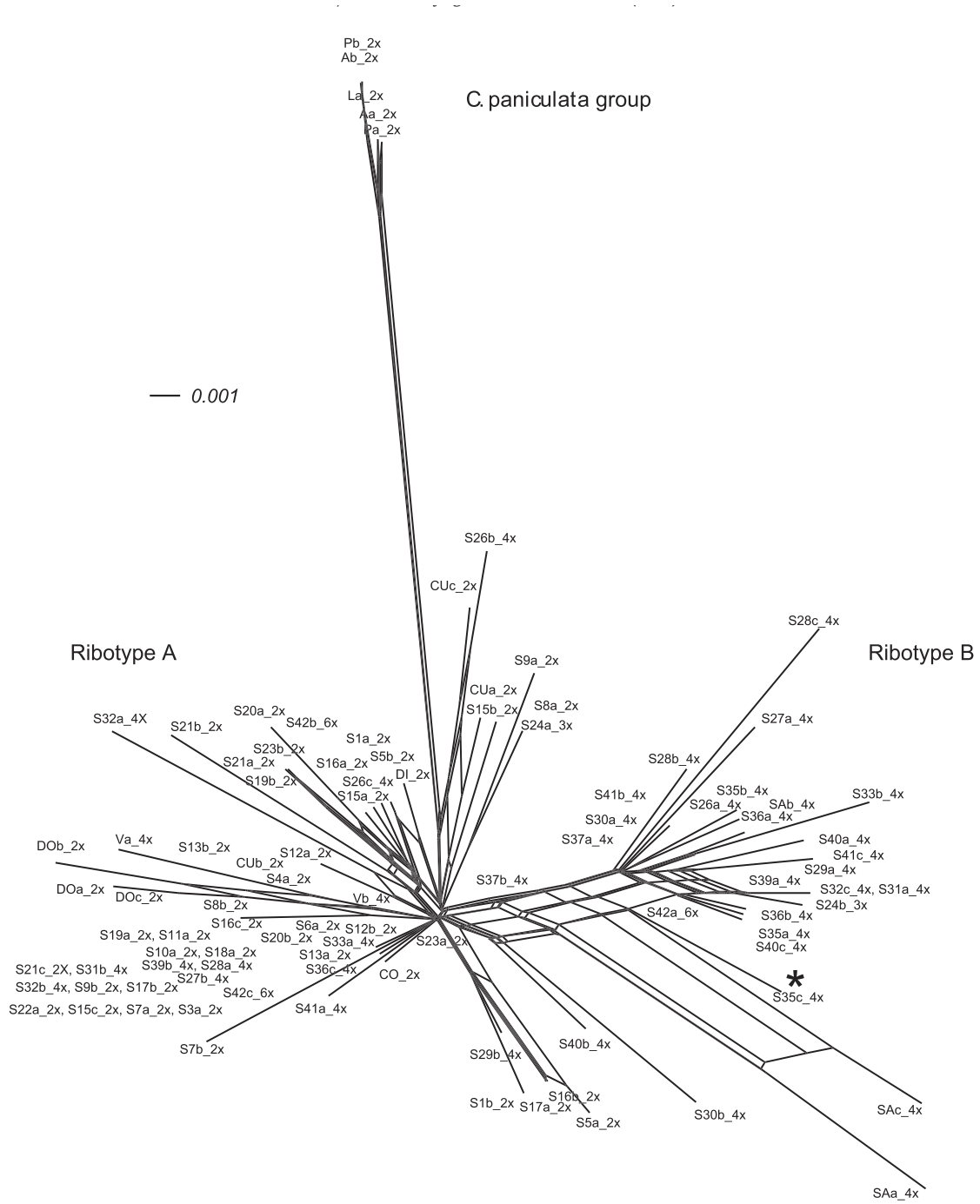

Fig. 2. NN split graphs based on uncorrected *p*-distances of the ITS matrix (non-informative and constant characters excluded). ITS copies A and B are tentatively indicated, and one putatively recombinant clone (S35c) is marked by asterisk. Plant codes are as follow: A – *C. aplolepa*; CO – *C. corymbosa*; CU – *C. cuneifolia*; DI – *C. diffusa*; DO – *C. donetzica*; L – *C. leucophea*; P – *C. paniculata*; SA – *C. sarandinakiae*; S1 to S42 – *C. stoebe*; V – *C. vankovii* (see Table 1).

Table 2
Diagnostic substitutions for A and B ribotypes found in the *Centaurea stoebe* group. Intermediate clones among both ribotypes are given below, the ploidy level and taxon's names are given in parentheses.

	77	199	577
Ribotype A	G	T	T
Ribotype B	A	C	G
A (S19b, S21a, S23b – all 2× *C. stoebe* s.str.)	G	C	T
B (SAa – *C. sarandinakiae* – 4×)	G	T	G
B (S35c – *C. stoebe* s.l. – 4×)	G	C	G

Neighbor-Net (NN) unrooted analysis confirmed the pattern obtained by the Bayesian and parsimony approach and revealed three main networks: one corresponding to the *C. paniculata* group, the second one to ribotype A, and the third one to ribotype B (Fig. 2). One ITS clone from a tetraploid individual (S35c) was placed in NN analysis between A and B networks together with two clones from *C. sarandinakiae* (SAa and SAc), suggesting its putative recombinant character (Fig. 2). In fact, the S35c clone exhibits G at site 77 as in ribotype A, while at sites 199 and 577 it has C and G respectively as in ribotype B (see above and Table 2).

3.2. Divergence time estimations based on ITS sequences

Mean genetic distance between the *Paniculata* clade and *Stoebe* clade using Tamura-Nei substitution model with gamma

Table 3
Split and diversification age estimations (in mya) calculated by a strict molecular clock approach and Bayesian analyses using relaxed molecular clock model (BEAST).

Method	Split of *C. paniculata* and *C. stoebe* clades	Origin of *C. stoebe* clade diversification	Origin of B ribotype diversification
Molecular clock (*Eupatorium* ITS substitution rate, 2.51×10^{-9})	8	–	–
Molecular clock (Madieae ITS substitution rate, 3×10^{-9})	9.6	–	
BEAST (95% HPD interval)	9.53 (7.73–11.37)	2.37 (1.41–5.21)	1.94 (0.62–2.46)

distribution was 0.024. When two slightly different calibrated substitution rates were used (see Section 2 and Table 3), the divergence time between the *C. paniculata* and the *stoebe* groups based on strict molecular clock approach was estimated to be c. 8 and 9.6 mya, respectively.

Bayesian estimations assuming relaxed ITS evolution over time provided slightly different age estimations than strict molecular clock method (cf. Table 3 and Fig. 3). Specifically, the split between the *C. paniculata* and C. *stoebe* clades was estimated to be c. 9.53 mya. The diversification of the *Stoebe* clade was estimated at 2.37 mya, and the origin of B ribotype diversification was estimated to be c. 1.94 mya (Table 3).

3.3. Haplotype diversity analyses

Combined *atpB–rbcL* and *trnT–trnL* sequences of 38 accessions resulted in 1303 bp long alignment. From 14 variable sites in total, six were parsimoniously informative. In addition to single nucleotide substitutions, 17 insertion–deletion polymorphisms were found, which were however excluded from subsequent analyses. Based on single nucleotide polymorphisms we constructed a haplotype network resulting in 12 different haplotypes (Fig. 4 and Table 1). No clear structure and largely shared cpDNA haplotype diversity in respect of analysed taxa and ploidies were found. The same pattern, i.e. no resolution was obtained using Maximum likelihood analysis (results not shown). In haplotype network the H1 and H2 haplotypes were the most frequent and included most of the diploids and tetraploids of the *C. stoebe* group, but in different proportions (Fig. 4). Three accessions belonging to three species from the *C. paniculata* group each belong to different haplotypes. While *C. paniculata* s.str. showed a unique haplotype (H9), two other species, *C. aplolepa* and *C. leucophaea*, shared their haplotypes (H2 and H3, respectively) with other accessions of *C. stoebe*. Three mutation steps present each twice on different branches may indicate frequent homoplasies (Fig. 4).

4. Discussion

4.1. Pattern of individual nrDNA polymorphism and hybridogeneous origin of tetraploid C. stoebe s.l.

Most of the accessions showed multiple ITS copies within each individual genome suggesting absence or very slow pace of concerted evolution. This pattern agrees with recent findings (reviewed in Bailey et al., 2003) that intra-individual polymorphism is much more frequent than previously thought (Baldwin et al., 1995). Generally, three main sources of intra-individual polymorphisms are recognized: hybridization (Kaplan and Fehrer, 2007; Záveská Drábková et al., 2009; Hirschegger et al., 2010; Jacob and Blattner, 2010; Šingliarová et al., 2011), ancestral polymorphism (Pamilo and Nei, 1988; Muir and Schlötterer, 2005) or presence of pseudogenes (Buckler and Holtsford, 1996; Kita and Ito, 2000; Mayol and Rosselló, 2001). In our case, the presence of pseudogene(s) can be ruled out, as cloned sequences showed no long indels, mutations in coding 5.8 S rDNA were extremely rare, the G–C content was in the range 56.1–57.4% and the ribotypes showed normal secondary structure (data not shown). Distinguishing between ancestral polymorphism and hybridization is more difficult. Ancestral polymorphism implies occurrence of all divergent ribotypes in the ancestral taxon and their various combinations in descendent heterozygous individuals. For this reason one may expect the presence of B ribotype also in some diploid *C. stoebe* s.str. accessions, which however was not the case and all diploid plants showed A ribotype (Figs. 1 and 2). Therefore, we propose a hybridization hypothesis to explain the co-occurrence of two divergent A and B ribotypes in almost all tetraploid individuals of *C. stoebe* s.l. One could argue, however, that such a pattern can reflect directional homogenization towards the A ribotype and loss of the B ribotype in diploid *C. stoebe* s.str. This hypothesis seems less plausible because a lack of concerted evolution and conservation of different ribotypes is the rule in the groups of *Centaurea* investigated to date (present data; Suárez-Santiago et al., 2007; Garcia-Jacas et al., 2009; Boršić et al., 2011; Hilpold, pers. comm.). Based on these results, we propose that one of the putative parental species of tetraploid *C. stoebe* s.l. is in fact diploid *C. stoebe* s.str., as both cytotypes share the A ribotype and are morphologically similar (Mráz et al., 2011). The donor of the B ribotype remains currently unknown. Although the B ribotype was found in tetraploid *C. sarandinakiae*, this species is morphologically clearly distinct from *C. stoebe* s.l. and very similar to *C. donetzica* and *C. vankovii*, which makes *C. sarandinakiae* as second parent very unlikely. Thus, the B progenitor could be either already extinct, as it was revealed in many polyploid complexes with exhausted sampling effort (Jacob and Blattner, 2010; Brokaw and Hufford, 2010) or the second parental species has not yet been sampled. Considering the fact that still new taxa from *Centaurea-Phalolepis* are being described each year mostly from East Mediterranean and Black Sea areas (Trigas et al., 2008; Doğan and Duran, 2009) we cannot exclude the possibility that the second parental taxon may still exist. The potential cradle of tetraploid *C. stoebe* s.l. may well be located in this area as SE Europe represents the diversity center of *Centaurea-Phalolepis* and the B ribotype was found in *C. sarandinakiae*, which is endemic to this region. Similarly, Ochsmann (2000) suggested that the tetraploid cytotype originated most likely in SE Europe, where the tetraploids are most frequent.

Molecular clock and Bayesian approach placed the split between the *C. paniculata* and *C. stoebe* clades within the late Miocene. On the other hand, diversification of the *Stoebe* group took place in a range from the late Miocene to the Pliocene (Bayesian estimation 1.41–5.21 mya, Fig. 3 and Table 3) and was probably favored by the climatic changes associated to the abrupt increase of the Mediterranean Sea level (Duggen et al., 2003). Interspecific hybridization(s) giving rise to tetraploid *C. stoebe* s.l. is expected to have occurred within the Pleistocene (Bayesian estimation 0.62–2.46 mya, Fig. 3 and Table 3), characterized by the alternation of glacial and interglacial cycles, pointing out its very young age.

Besides the tetraploid cytotype, co-occurrence of A and B ribotypes was found in one triploid individual collected in a mixed-ploidy population, and in one hexaploid plant found in one tetraploid population from the introduced range (Mráz et al., 2011). Since this triploid is from a mixed-ploidy site (Mráz et al., unpubl.),

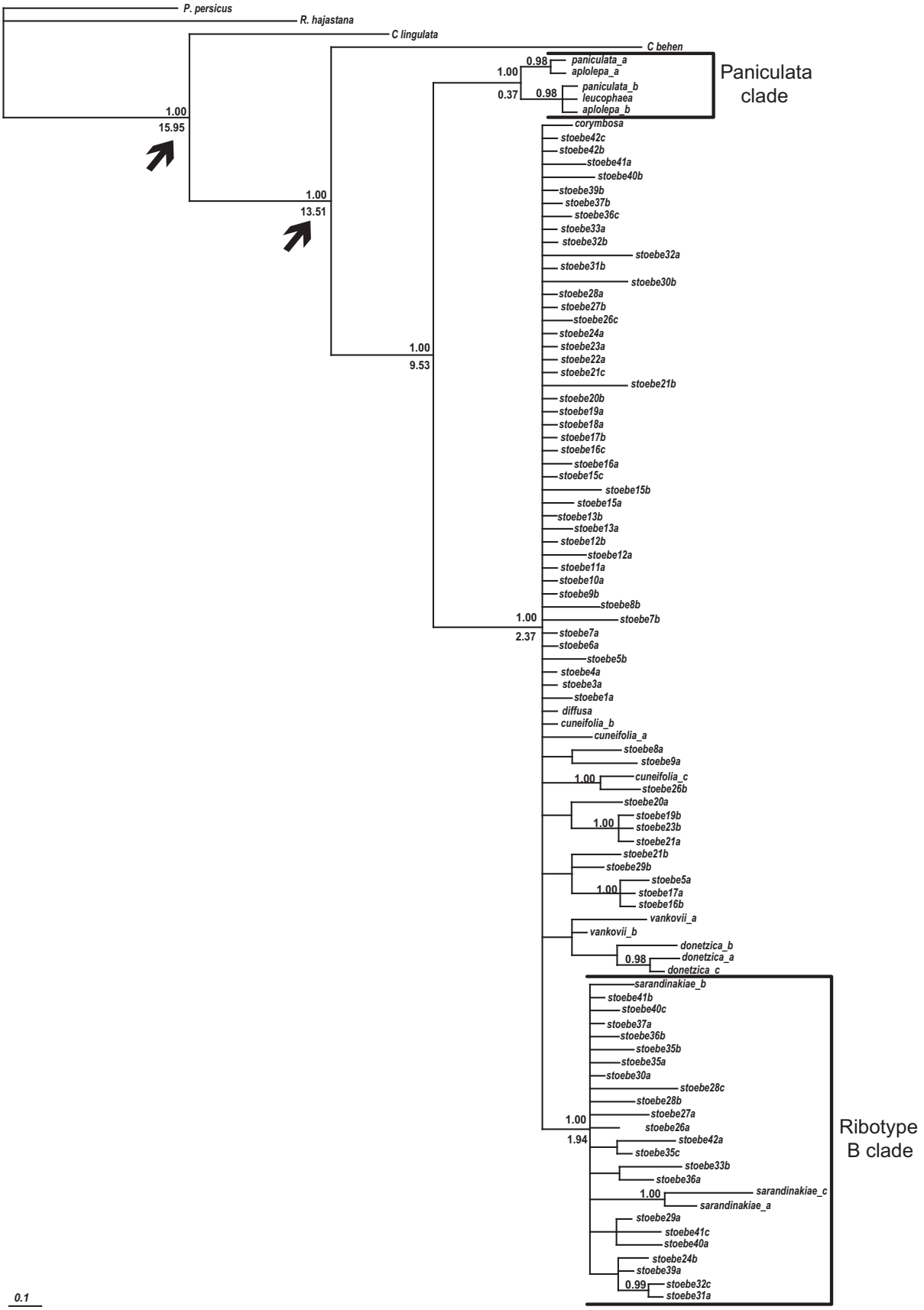

Fig. 3. Bayesian majority rule consensus tree from the ITS datation dataset. Numbers above branches indicate Bayesian-credibility values (PP) and numbers under branches indicate estimated ages of the main supported clades using BEAST. Two calibration data points labeled by arrows were used (see Section 2).

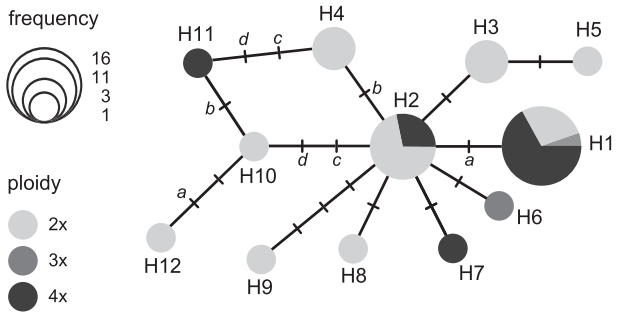

Fig. 4. Haplotype network of twelve haplotypes from 38 accessions of the *Centaurea stoebe* and *C. paniculata* groups based on combined *trn*T–*trn*L and *rbc*L–*atp*B sequences. Three different ploidies are marked by different shading, one hexaploid accession of *C. stoebe* s.l. (URS3) belonging to H1 haplotype is not distinguished. Putatively homoplasious mutations are labeled by italic lowercase letters.

its hybridogenous origin between a 2× and 4× plant is most likely. Such a scenario is corroborated not only by the ITS pattern, but also by microsatelite analyses (Mráz et al., unpubl.). The origin of the single hexaploid plant found in the introduced range in Northern America has been explained through fusion of reduced ($n = 18$) and unreduced gametes of tetraploid plants ($n = 36$, cf. Mráz et al., 2011) and therefore the sharing of A and B ribotypes in this hexaploid is logic.

The allotetraploid origin of *C. stoebe* s.l. further questions the single species concept proposed by Španiel et al. (2008) and rather supports the taxonomic recognition of the tetraploid cytotype as a separate taxon (Ochsmann, 2000; Mráz et al., 2011). Our molecular data are further supported by strong reproductive barriers between both cytotypes (Mráz et al., unpubl.), differences in morphology and life-cycle (Mráz et al., 2011) and the slightly lower homoploid genome size found in tetraploids (Bancheva and Greilhuber, 2006; Mráz and Keller, unpubl.).

Although the second parental species of tetraploid *C. stoebe* s.l. remains so far unknown, we assume that hybridization stimulated the phenotypic and life-history change between both cytotypes, as interspecific hybridization is considered a prominent and instant mechanism of creating new variation (Arnold, 1992; Rieseberg, 1997). The change from the annual/biennial monocarpic life cycle to a perennial one could explain the colonization success of polycarpic tetraploids not only in the introduced range, but also in Central and Western Europe, where the massive spread of 4× cytotype has been recently recorded (e.g. Ochsmann, 2000). Indeed, a perennial life cycle assuring greater persistence and extended seed production may constitute a more efficient strategy for colonizing mesophilous climates like those found in Central Europe. This is in a striking contrast with the conditions in the Mediterranean region, where most successful colonizers among Cardueae are biennial monocarps (Garcia-Jacas et al., 2008). Although polyploidy is often assumed to be an important trait explaining invasion or colonization success (e.g. Ehrendorfer, 1980; Stebbins, 1985; Verlaque et al., 2002; Brochmann et al., 2004; Küster et al., 2008; Pyšek et al., 2009; Pandit et al., 2011), in many cases it is only the consequence of interspecific hybridization (Paun et al., 2009). In this view, interspecific hybridization seems to be a more efficient speciation mechanism as compared to polyploidization *per se*, which in turn acts mostly as a "stabilizing" mechanism of the breeding behavior in newly created hybrids and hybridogeneous species (Grant, 1981). It is thus necessary to distinguish between auto- and allopolyploids in studies focusing on causes of plant invasiveness, as hybridization could be of greater importance for the formation of "evolutionary novelty" leading, in our case, to better colonization and persistence than polyploidization alone.

4.2. ITS variation and phylogenetic relationships within and between the C. stoebe and C. paniculata groups

The A ribotype was the most common ribotype and was found in all taxa and cytotypes of *C. stoebe*, including the taxa which were considered by Ochsmann (2000) to be morphologically either closely related to (e.g. *C. reichenbachii*, *C. triniifolia*) or distinct from *C. stoebe* (*C. corymbosa*, *C. diffusa*, *C. vallesiaca*). *Centaurea reichenbachii* and *C. triniifolia* are morphologically indistinguishable from *C. stoebe* and accordingly, they do not deserve taxonomic recognition (cf. Mráz et al., 2011; Mráz, unpubl.). On the other hand, *C. corymbosa*, *C. cuneifolia*, *C. diffusa* and *C. vallesiaca* are phenotypically well differentiated in spite of their similar ITS sequences. Such pattern may indicate very recent diversification of this group probably associated with range isolation as in the case of *C. corymbosa* and *C. vallesiaca*. Both of these species are endemics of small regions outside of the continuous *C. stoebe* range. Furthermore, adaptation to specific habitats, such as crevices of calcareous cliffs in the case of the Mediterranean *C. corymbosa* (Colas et al., 1997), or extremely dry steppes in the Black Sea region in *C. diffusa*, could further accelerate morphological differentiation of these taxa.

Our study confirms the results of Ochsmann (2000) and Suárez-Santiago et al. (2007) that the *C. paniculata* group is well separated from the *C. stoebe* group. The distinct position of the *C. paniculata* group is furthermore supported by morphological differences (Ochsmann, 2000), and by c. 1.5 higher homoploid genome size at diploid level in the *C. paniculata* group (Mráz, unpubl.). These finding thus challenge the suggested relationships of the *C. paniculata* group and the bulk of section *Centaurea* (syn. *Acrolophus*, cf. Wagenitz and Hellwig, 1996) and relate it rather to the West Mediterranean section *Willkommia*. The estimated divergence time between the *C. stoebe* and *C. paniculata* groups (c. 8–11 mya) largely overlaps with the Tortonian stage of Miocene characterized by northwards shift of bioms due to increased precipitations and temperature (Pound et al., 2011). This phenomenon could contribute to the fragmentation of the range of common ancestors, and geographically separated populations could diverge.

4.3. Shared plastid DNA diversity

In contrast to ITS sequences, two combined cpDNA loci did not reveal any clear structure with respect to the taxonomic position or ploidy level of the studied accessions. Sharing of different cpDNA haplotypes could be explained by (i) interspecific gene flow, (ii) ancestral polymorphism and (iii) an independent origin of haplotypes. Based on present and still unpublished data from more than 900 accessions (Treier et al., unpubl.) it seems that all three mutually non-exclusive explanations may act in concert. Firstly, extensive gene flow between closely related species of *Centaurea* is a largely accepted fact (see Section 1) and is obvious also in the sect. *Centaurea* (Ochsmann, 2000). Ancestral cpDNA polymorphism and slow mutational rate of this marker could further, at least partly, explain the observed shared cpDNA haplotypes between the phylogenetically divergent *C. stoebe* and *C. paniculata* taxa. Finally, occurrences of homoplasious mutations could further blur the cpDNA pattern (see Fig. 4). Our results thus suggest that cpDNA cannot be used to infer species relationships in *Centaurea* and thus confirm findings from other *Centaurea* sections and groups (Font et al., 2009; Garcia-Jacas et al., 2009; Löser et al., 2009).

4.4. Conclusions

Cloning of nrDNA revealed evidence for a hybridogeneous origin of tetraploid *Centaurea stoebe* s.l. Together with differences in morphology, life cycle and homoploid genome size, allopolyploidization provides a further argument for taxonomic recognition of

the tetraploid cytotype as a different species. Although we do not know the second parental taxon, hybridization could have triggered important changes in the phenotype and life-cycle of the newly formed allotetraploid taxon. Such changes could contribute to better colonization abilities and thus invasion success of the tetraploid cytotype as compared to one of its progenitors – the diploid *C. stoebe* s.str. Our data furthermore indicate very recent diversification of the highly variable *C. stoebe* group, where some morphologically distinct taxa diverged in geographical allopatry. On the basis of cloned ITS sequences and in congruence with other studies (Ochsmann, 2000; Garcia-Jacas et al., 2006; Suárez-Santiago et al., 2007), the *C. paniculata* group is clearly distinct from most of the rest of species of sect. *Centaurea* (syn. *Acrolophus*), but closely related to sect. *Willkommia*, which consists of West Mediterranean taxa.

Acknowledgments

We thank A-C. Cossy for help in the molecular lab, all collectors for providing seed material (see Table 1) and two anonymous reviewers for valuable comments. The work was financially supported through the National Centre of Competence in Research (NCCR) Plant Survival, a research programme of the Swiss National Science Foundation (to HMS), the Spanish Ministry of Science & Innovation (Project CGL2007-60781/BOS) and the Autonomous Government of Catalunya (''Ajuts a Grups de Recerca Consolidats'' 2009/SGR/00439) (to NGJ, AS and LB). P. Mráz and E. Gex-Fabry thank the University of Fribourg (mobility grant) for funding of their stays in Barcelona and Llorenç Sáez (Autonomous University of Barcelona) for hosting EGF.

References

Abbott, R.J., Hegarty, M.J., Hiscock, S.J., Brennan, A.C., 2010. Homoploid hybrid speciation in action. Taxon 59, 1375–1386.
Ainouche, M.L., Baumel, A., Salmon, A., Yannic, G., 2004. Hybridization, polyploidy and speciation in *Spartina* (Poaceae). New Phytol. 161, 161–172.
Álvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29, 417–434.
Anderson, E., Stebbins, G.L., 1954. Hybridization as an evolutionary stimulus. Evolution 8, 378–388.
Andreasen, K., Baldwin, B.G., 2003. Nuclear ribosomal DNA sequence polymorphism and hybridization in checker mallows (*Sidalcea*, Malvaceae). Mol. Phylogenet. Evol. 29, 563–581.
Arnold, M.J., 1992. Natural hybridization as an evolutionary process. Ann. Rev. Ecol. Syst. 23, 237–261.
Bailey, C.D., Carr, T.G., Harris, S.A., Hughes, C.E., 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol. Phylogenet. Evol. 29, 435–455.
Baldwin, B.G., Sanderson, M.J., 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). proc. Natl. Acad. Sci. USA 95, 9402–9406.
Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, C.S., Donoghue, M.J., 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence of angiosperm phylogeny. Ann. Miss. Bot. Garden 82, 247–277.
Bancheva, S., Greilhuber, J., 2006. Genome size in Bulgarian *Centaurea* s.l. (Asteraceae). Plant Syst. Evol. 257, 95–117.
Bandelt, H-J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.
Baumel, A., Ainouche, M.L., Bayer, R.J., Ainouche, A.K., Misset, M.T., 2002. Molecular phylogeny of hybridizing species from the genus *Spartina* Schreb (Poaceae). Mol. Phylogenet. Evol. 22, 303–314.
Blair, A.C., Huffbauer, R.A., 2010. Hybridization and invasion: one of North America's most devastating invasive plants shows evidence for a history of interspecific hybridization. Evol. Appl. 3, 40–51.
Blakeslee, A.F., 1941. Effect of induced polyploidy in plants. Am. Nat. 75, 117–135.
Boggs, K.W., Story, J.M., 1987. The population age structure of spotted knapweed (*Centaurea maculosa*) in Montana. Weed Sci. 35, 194–198.
Boršić, I., Susanna, A., Bancheva, S., Garcia-Jacas, N., 2011. *Centaurea* sect. *Cyanus*: nuclear phylogeny, biogeography and life-form evolution. Int. J. Plant Sci. 172, 238–249.
Brochmann, C., Brysting, A.K., Alsos, I.G., Borgen, L., Grundt, H.H., Scheen, A.-C., Elven, R., 2004. Polyploidy in arctic plants. Biol. J. Linn. Soc. 82, 521–536.
Brokaw, J.M., Hufford, L., 2010. Origins and introgression of polyploid species in *Mentzelia* section *Trachyphytum* (Loasaceae). Am. J. Bot. 97, 1457–1473.
Bryant, D., Moulton, V., 2004. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. Mol. Biol. Evol. 21, 255–265.
Buckler, E.S., Holtsford, T.P., 1996. *Zea* ribosomal repeat evolution and substitution patterns. Mol. Biol. Evol. 13, 623–632.
Burke, J.M., Arnold, M.L., 2001. Genetics and the fitness of hybrids. Annu. Rev. Genet. 35, 31–52.
Campbell, C.S., Wojciechowski, M.F., Baldwin, B.G., Alice, L.A., Donoghue, M.J., 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). Mol. Biol. Evol. 14, 81–90.
Chiang, T.Y., Schaal, B.A., Peng, C.I., 1998. Universal primers for amplification and sequencing a noncoding spacer between the *atpB* and *rbcL* genes of chloroplast DNA. Bot. Bull. Acad. Sin. (Taipei) 39, 245–250.
Cline, J., Braman, J.C., Hogrefe, H.H., 1996. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. Nucl. Acids Res. 24, 3546–3551.
Comai, L., 2005. The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 6, 836–846.
Colas, B., Olivieri, I., Riba, M., 1997. *Centaurea corymbosa*, a cliff dwelling species tottering on the brink of extinction: a demographical and genetic study. Proc. Natl. Acad. Sci. USA 94, 3471–3476.
Doğan, B., Duran, A., 2009. *Centaurea serpentinica* sp. nov. (Asteraceae) from central and south Anatolia transition zone, Turkey. Nord. J. Bot. 27, 319–323.
Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15.
Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Duggen, S., Hoernle, K., Van Den Bogaard, P., Rupke, L., Morgan, J.P., 2003. Deep roots of the Messinian salinity crisis. Nature 422, 602–606.
Ehrendorfer, F., 1980. Polyploidy and distribution. In: Lewis, W.H. (Ed.), Polyploidy—Biological Relevance. Plenum Press, New York, pp. 45–60.
Ellstrand, N.C., Schierenbeck, K.A., 2000. Hybridization as a stimulus for the evolution of the invasiveness in plants. Proc. Natl. Acad. Sci. USA 97, 7043–7050.
Fehrer, J., Krak, K., Chrtek, J., 2009. Intra-individual polymorphism in diploid and apomictic polyploid hawkweeds (*Hieracium*, Lactuceae, Asteraceae): disentangling phylogenetic signal, reticulation and noise. BMC Evol. Biol. 9, 239.
Fernández Casas, J., Susanna, A., 1986. Monografía de la sección *Chamaecyanus* Willkomm del género *Centaurea* L. Treb. Inst. Bot. Barcelona 10, 1–174.
Font, M., Garcia-Jacas, N., Vilatersana, R., Roquet, C., Susanna, A., 2009. Evolution and biogeography of *Centaurea* section *Acrocentron* inferred from nuclear and plastid DNA sequence analyses. Ann. Bot. 103, 985–997.
Garcia-Jacas, N., Galbany-Casals, M., Romashchenko, K., Susanna, A., 2008. On the conflicting generic delineation in the *Onopordum* group (Compositae, Cardueae-Carduinae): a combined nuclear and plastid molecular approach. Aust. Syst. Bot. 21, 301–311.
Garcia-Jacas, N., Soltis, P.S., Font, M., Soltis, D.E., Vilatersana, R., Susanna, A., 2009. The polyploidy series of *Centaurea toletana*: glacial migrations and introgression revealed by nrDNA and cpDNA sequence analyses. Mol. Phylogenet. Evol. 52, 377–394.
Garcia-Jacas, N., Susanna, A., 1994. *Centaurea prolongi* and *Centaurea crocata* in Portugal: an old confusion. Nord. J. Bot. 14, 31–38.
Garcia-Jacas, N., Uysal, T., Romaschenko, K., Suárez-Santiago, V.N., Ertuğrul, K., Susanna, A., 2006. *Centaurea* revisited: a molecular survey of the *Jacea* group. Ann. Bot. 98, 741–753.
Grant, V., 1981. Plant Speciation. Columbia University Press, New York, USA.
Gray, A.J., Benham, P.E.M., Raybould, A.R., 1990. *Spartina anglica* – the evolutionary and ecological background. In: Gray, A.J., Benham, P.E.M. (Eds.), *Spartina anglica* – A Research Review. Institute of Terrestrial Ecology Research Publication No. 2, HMSO, London, UK, pp. 5–10.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignement editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Hardy, O.J., Vanderhoeven, S., de Loose, M., Meerts, P., 2000. Ecological, morphological and allozymic differentiaton between diploid and tetraploid knapweeds (*Centaurea jacea*) from a contact zone in the Belgian Ardennes. New Phytol. 146, 281–290.
Hellwig, F.H., 2004. Centaureinae (Asteraceae) in the Mediterranean – history of ecogeographical radiation. Plant Syst. Evol. 246, 137–162.
Henery, M.L., Bowman, G., Mráz, P., Treier, U.A., Gex-Fabry, E., Schaffner, U., Müller-Schärer, H., 2010. Evidence for a combination of pre-adapted traits and rapid adaptive change in the invasive plant *Centaurea stoebe*. J. Ecol. 98, 800–813.
Hirschegger, P., Jakše, J., Trontel, j.P., Bohanec, B., 2010. Origins of *Allium ampeloprasum* horticultural groups and a molecular phylogeny of the section *Allium* (*Allium*: Alliaceae). Mol. Phylogenet. Evol. 54, 488–497.
Ho, S.Y.M., 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. J. Avian Biol. 38, 409–414.
Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenenetic trees. Bioinformatics 17, 754–755.
Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267.
Jacob, S.S., Blattner, F.R., 2010. Two extinct diploid progenitors were involved in allopolyploid formation in the *Hordeum murinum* (Poaceae: Triticaceae) taxon complex. Mol. Phylogenet. Evol. 55, 650–659.

Kaplan, Z., Fehrer, J., 2007. Molecular evidence for a natural primary triple hybrid in plants revealed from direct sequencing. Ann. Bot. 99, 1213–1222.
Kay, M.K., Whittall, J.B., Hodges, S.A., 2006. A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects. BMC Evol. Biol. 6, 36.
Kim, S.T., Sultan, S.E., Donoghue, M.J., 2008. Allopolyploid speciation in *Persicaria* (Polygonaceae): insight from a low-copy nuclear region. Proc. Natl. Acad. Sci. USA 105, 12370–12375.
Kita, Y., Ito, M., 2000. Nuclear ribosomal ITS sequences and phylogeny in East Asian *Aconitum* subgenus *Aconitum* (Ranunculaceae) with special reference to extensive polymorphism in individual plants. Plant Syst. Evol. 225, 1–13.
Koutecký, P., 2007. Morphological and ploidy level variation of *Centaurea phrygia* agg. (Asteraceae) in the Czech Republic, Slovakia and Ukraine. Folia Geobot. 42, 77–102.
Kummer, C., 1977. Untersuchungen von natürlichen Bastarden zwischen *Centaurea cephalariifolia* Willkomm und *Centaurea ornata* Willdenow. Mitt. Bot. Staatssamml. München 13, 129–202.
Küster, E.C., Kühn, I., Bruelheide, H., Klotz, S., 2008. Trait interactions help explain invasion success in the German flora. J. Ecol. 96, 860–868.
Löser, C.J., Akaydin, G., Hellwig, F.H., 2009. Incomplete lineage sorting among annuals and perennials in the genus *Cyanus*. Book of abstracts, systematics 2009. In: 7th Biennial Conference of the Systematics Association, Leiden, August 10–14.
Maherali, H., Walden, A.E., Husband, B.C., 2009. Genome duplication and the evolution of physiological responses to water stress. New Phytol. 184, 721–731.
Mayol, M., Rosselló, J.A., 2001. Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. Mol. Phylogenet. Evol. 19, 167–176.
Mráz, P., Bourchier, R.S., Treier, U.A., Schaffner, U., Müller-Schärer, H., 2011. Polyploidy in phenotypic space and invasion context: a morphometric study of *Centaurea stoebe* s.l. Int. J. Plant Sci. 172, 386–402.
Muir, G., Schlötterer, C., 2005. Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). Mol. Ecol. 14, 549–561.
Müller, H., 1989. Growth pattern of diploid and tetraploid spotted knapweed, *Centaurea maculosa* Lam. (Compositae), and effects of the root-mining moth *Agapeta zoegana* (L) (Lep.: Cochylidae). Weed Res. 29, 103–111.
Müntzing, A., 1936. The evolutionary significance of autopolyploidy. Hereditas 21, 263–378.
Nylander, J.A.A., 2004. MrModeltest v. 2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University.
Ochsmann, J., 2000. Morphologische und molekularsystematische Untersuchungen an der *Centaurea stoebe* L. – Gruppe (Asteraceae–Cardueae) in Europa. Diss. Bot., 324.
Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5, 568–583.
Pandit, M.K., Pocock, Michael J.O., Kunin, W.E., 2011. Ploidy influences rarity and invasiveness in plants. J. Ecol. 99, 1108–1115.
Paun, O., Forest, F., Fay, M.F., Chase, M.W., 2009. Hybrid speciation in angiosperms: parental divergence drives ploidy. New Phytol. 182, 507–518.
Pisanu, S., Mameli, G., Farris, E., Binelli, G., Filigheddu, R., 2011. A natural homoploid hybrid between *Centaurea horrida* and *Centaurea filiformis* (Asteraceae) as revealed by morphological and genetic traits. Folia Geobot. 46, 69–86.
Popp, M., Oxelman, B., 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. Mol. Phylogenet. Evol. 20, 474–481.
Posada, D., 2008. jModeltest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Pound, M.J., Haywood, A.M., Salzmann, U., Riding, J.B., Lunt, D.J., Hunter, S.J., 2011. A tortonian (late Miocene, 11.61–7.25 Ma) global vegetation reconstruction. Palaeogeogr. Palaeoclim. Palaeoecol. 300, 29–45.
Pyšek, P., Jarošík, V., Pergl, J., Randall, R., Chytrý, M., Kuhn, I., Tichý, L., Danihelka, J., Chrtek, J.J., Sádlo, J., 2009. The global invasion success of Central European plants is related to distribution characteristics in their native range and species traits. Div. Distr. 15, 891–903.
Rieseberg, L.H., 1997. Hybrid origin of plant species. Ann. Rev. Ecol. Syst. 28, 359–389.
Rieseberg, L.H., Soltis, D.E., 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plant 5, 65–84.
Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Sang, T., Crawford, D.J., Stuessy, T.F., 1995. Documentation of reticulate evolution in peonies (*Peonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA – implications for biogeography and concerted evolution. Proc. Natl. Acad. Sci. USA 92, 6813–6817.
Schmidt, G.J., Schilling, E.E., 2000. Phylogeny and biogeography of *Eupatorium* (Asteraceae: Eupatorieae) based on nuclear ITS sequence data. Am. J. Bot. 87, 716–726.
Sheley, R.L., Jacobs, J.S., Carpinelli, M.F., 1998. Distribution, biology and management of diffuse knapweed (*Centaurea diffusa*) and spotted knapweed (*Centaurea maculosa*). Weed Technol. 12, 353–362.
Šingliarová, B., Hodálová, I., Mráz, P., 2011. Biosystematic study of the diploid–polyploid *Pilosella alpicola* group with variation in breeding system: patterns and processes. Taxon 60, 450–470.
Soltis, D.E., Kuzoff, R.K., 1995. Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). Evolution 49, 727–742.
Soltis, P.S., Soltis, D.E., 1991. Multiple origins of the allotetraploid *Tragopogon mirus* (Compositae): rDNA evidence. Syst. Bot. 16, 407–413.
Španiel, S., Marhold, K., Hodálová, I., Lihová, J., 2008. Diploid and tetraploid cytotypes of *Centaurea stoebe* (Asteraceae) in Central Europe: morphological differentiation and cytotype distribution patterns. Folia Geobot. 43, 131–158.
Stebbins, G.L., 1950. Variation and Evolution in Plants. Columbia Univesrity Press, New York, USA.
Stebbins, G.L., 1985. Polyploidy, hybridization, and the invasion of the new habitats. Ann. Miss. Bot. Gard. 72, 824–832.
Stelkens, R.B., Schmid, C., Selz, O., Seehausen, O., 2009. Phenotypic novelty in experimental hybrids is predicted by the genetic distance between species of cichlid fish. BMC Evol. Biol. 9, 283.
Sun, Y., Skinner, D.Z., Liang, G.H., Hulbert, S.H., 1994. Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor. Appl. Genet. 89, 26–32.
Story, J.M., Smith, L., Good, W.R., 2001. Relationship among growth attributes of spotted knapweed (*Centaurea maculosa*) in western Montana. Weed Technol. 15, 750–761.
Suárez-Santiago, V.N., Salinas, M.J., Garcia-Jacas, N., Soltis, P.S., Soltis, D.E., Blanca, G., 2007. Reticulate evolution in the *Acrolophus* subgroup (*Centaurea* L, Compositae) from the western Mediterranean: origin and diversification of section *Willkommia* Blanca. Mol. Phylogenet. Evol. 43, 156–172.
Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA 4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
Thompson, J.D., 1991. The biology of an invasive plant – what makes *Spartina anglica* so successful. Bioscience 41, 391–401.
Treier, U.A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T., Müller-Schärer, H., 2009. Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. Ecology 90, 1366–1377.
Trigas, P., Constantinidis, T., Touloumenidou, T., 2008. A new hexaploid species of *Centaurea* section *Acrolophus* (Asteraceae) from Evvia Island, Greece. Bot. J. Linn. Soc. 158, 762–774.
Vanderhoeven, S., Brown, C.S., Tepolt, C.K., Tsutsui, N.D., Vanparys, V., Atkinson, S., Mahy, G., Monty, A., 2010. Linking concepts in the ecology and evolution of invasive plants: network analysis shows what has been most studied and identifies knowledge gaps. Evol. Appl. 3, 193–202.
Verlaque, R., Aboucaya, A., Frindler, A., 2002. Invasive alien flora of France: ecology, life-forms and polyploidy. Bot. Helv. 112, 121–136.
Vilatersana, R., Brysting, A.K., Brochmann, C., 2007. Molecular evidence for hybrid origins of the invasive polyploids *Carthamus creticus* and *C. Turkestanicus* (Cardueae, Asteraceae). Mol. Phylogenet. Evol. 44, 610–621.
Wagenitz, G., Hellwig, F., 1996. Evolution of characters and phylogeny of Centaureinae. In: Hind, D.J.N., Beentje, H.J. (Eds.), Compositae – Systematics. Royal Botanical Gardens, Kew, UK, pp. 491–510.
Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.
Záveská Drábková, L., Kirschner, J., Štěpánek, J., Záveský, L., Vlček, Č., 2009. Analysis of nrDNA polymorphism in closely related diploid sexual, tetraploid sexual and polyploid agamospermous species. Plant Syst. Evol. 278, 67–85.