BABEȘ-BOLYAI UNIVERSITY CLUJ - NAPOCA

Faculty of Biology and Geology

Summary of PhD Thesis

Molecular phylogenetic studies on the European species of the genus *Astragalus* L. section *Dissitiflori*

Scientific supervisor

Prof. Dr. Nicolae Dragoș

PhD student

László Bartha

Cluj - Napoca

2012

Table of contents

Introduction

Astragalus is the largest genus (cca. 2500 spp.) of flowering plants. Beside the huge number of species, the diversity of genus is reflected also by many local endemics throughout the World, or rampant speciation e.g. in the Andes or Central Asia. The Romanian flora, for example, includes three endemic species: *A. peterfii* Jáv., *A. roemeri* Simonk., and *A. pseudopurpureus* Guşul. They are emblematic flagship species of Romanian nature conservation all of them having a restricted geographic distribution. The first one grows in steppe-like habitats of the Transylvanian Lowland, whereas the latter two are confined to mountain habitats of the Eastern Carpathians.

Exploring the relationships of such species is the obligation of current botanical researches. With the aid of DNA sequencing and phylogenetic tree reconstruction methods, fulfillment of such needs seems relatively straightforward. Generally, the more closely related the species are, the more similar are their sequences. Several methods commonly used for phylogenetic inference are based on this principle. An important issue these type of studies usually face is adequate taxon sampling. Which taxa to include among those cca. 2500? Previous taxonomic treatments based on morphological investigations or earlier phylogenetic studies on the genus can serve as guidelines for this purpose.

The interest of the author in plant phylogenetic inference had been raised during his M.Sc. studies and was heavily influenced by researchers from the Babeş-Bolyai University and the University of Debrecen. At the beginning of his PhD studies investigations aimed at reconstructing the molecular phylogenetic relationship of the narrow endemic *Astragalus peterfii*. Later, the author expanded the taxon sampling to whole Europe as far as that taxonomic section is considered to which *A. peterfii* belonged. The ultimate goal of his Ph.D. studies became to reconstruct this groups' phylogeny in Europe using state-of-the-art research methods. In the thesis the most important results of this three-year long research are presented.

Theoretical background

Inference of species relationships using DNA sequences

For more than two decades now, DNA sequences have been the main source of information for inference of species relatedness at almost all taxonomic levels and are going to be in the future (Baldwin, 1992; Shaw *et al.*, 2005). An antecedent of this was the recognition that DNA sequence variation (between taxa) was most likely to reflect more reliable relationships of species as compared with morphological characters which are prone to be highly homoplasious due to morphological convergence, and therefore to flaws in systematics (Downie *et al.*, 2000; Richardson *et al.*, 2000). Reliability is especially true for non-coding phylogenetic markers that are generally free from selective pressure, therefore are likely to depict a neutral picture of species' history (Kolarčik *et al.*, 2010). Two additional key factors have also contributed to the popularity of sequence based molecular phylogenetic inference: generation of sequences in a routine manner and advances in computational techniques and algorithms.

Phylogenies have direct implications to phylogenetic systematics: a clade of a phylogenetic tree is a (monophyletic) unit to which a taxonomic unit can be assigned if the clade in question is also supported by a morphological character or a combination of morphological characters (Chatrou *et al.*, 2012). An important issue to this approach is the (statistical) confidence of a certain clade. Only statistically (at least) well supported, robust clades should be considered. Moreover, the support, resolution and taxa content of any clade can vary in face of new data (sequences and/or more taxa) or when different phylogenetic tree reconstruction methods are applied. Robustness, therefore, is an important feature in phylogenetic systematics: taxonomic units are usually drawn from congruent, stable, statistically and morphologically supported clades.

Within angiosperms nucleotide substitution rates of plastid genomes usually exceed three times, while that of nuclear genomes 12–16 times of the mitochondrial genome (Weng *et al.*, 2012)), and references therein). Not surprisingly, in stark contrast with animals, mitochondrial sequences have gained less importance in plant phylogenetics.

Plastid DNA markers, even if they are usually less variable than the nuclear ones, present a sum of advantages when compared to nuclear regions. They usually reflect the classical (bifurcating) way of speciation on a phylogenetic tree due to their single-copy feature within an organism, which also ensures the lack of recombination. Another advantage is the accessibility to universal primers (Xu *et al.*, 2012). The above accounts make these types of markers ideal for e.g. testing monophyly (especially above the species level) even if they reflect the species' history only from the perspective of the maternal lineage (since they are usually inherited maternally). Plastid

DNA phylogenies, however, can often have polytomies in analyses of datasets that lack sufficient levels of sequence variation. Such cases can be explained by the recent and/or rapid diversification of the plant-group in question or by the use of inappropriate (i.e. relatively slowly evolving) markers.

Phylogenetic relationships, especially those towards the tips of a phylogenetic tree (at the species level), cannot be fully uncovered only by the use of uniparentally inherited markers when hybridisation, introgression, gene transfer (and other processes of similar output) shaped speciation. Hybridisation and allopolyploidisation (hybridisation followed by polyploidisation), for example, are recognised as major forces driving plant evolution and speciation (Grant, 1981; Soltis and Soltis, 2009). These processes can be identified by phylogenies inferred from biparentally inherited nuclear markers if copies of a nuclear region originating from different putative parental species are retrieved from a hybrid organism and those copies were sufficiently diverged prior to hybridisation. If either of such copies is homogenised towards another parental copy-type – a phenomena commonly occurring in case of the internal transcribed spacer region of nuclear ribosomal DNA (Álvarez and Wendel, 2003) – inconsistencies between the nuclear and plastid DNA phylogenies may still serve as evidence for reticulate (hybrid) speciation, or eventually allopolyploidisation. Incongruencies between nuclear and plastid DNA phylogenies (or paralogy of any particular nuclear loci), however, can also be caused by incomplete lineage sorting and recombination (Rieseberg and Brunsfeld, 1992; Xu *et al.*, 2012). Discriminating reticulation from incomplete lineage sorting (i.e. retention of ancestral polymorphism) is often difficult and represents a challenging task for evolutionary biologists (Wendel and Doyle, 1998; Willyard *et al.*, 2009).

In spite of the rapidly growing popularity of the low/single-copy nuclear genes, the internal transcribed spacer (ITS) region of the nuclear ribosomal DNA (hereafter referred to as nrITS region) is still one of the most widely used plant phylogenetic markers. It comprises the ITS1 and ITS2 regions linked by the 5.8S rRNA gene (Calonje *et al.*, 2008). ITS1 and ITS2 are transcribed during transcription but degraded after, thus they can be considered as neutral phylogenetic markers. Due to the high copy number within the genome, the relative shortness (i.e. the length is usually only several hundred base pairs) and existing universal primers, amplification of nrITS is relatively straightforward even from partially degraded DNA samples (e.g. extraction from herbarium material). However, repeats of this marker sometimes are not fully homogenised or do not track both parent's genomes in hybrids (Zimmer and Wen, 2012), and references therein) due to the already mentioned concerted evolution, making it a less promising tool as compared with low-copy nuclear genes.

General overview of the genus *Astragalus*

Astragalus: an evolutionary success story

Astragalus L. is the most species rich plant genus among angiosperms with approximately 2,500 species (Lock and Schrire, 2005). Species of this genus are non-climbing herbs, mostly perennials and to a lesser extent annuals. *Astragalus* species are present on all continents, excluding Australia, and occupy mostly the cool, arid, continental regions of the northern hemisphere and South America (Lock and Schrire, 2005). The genus reaches its greatest diversity in southwest and central Asia (ca. 1500 spp.), the Sino-Himalayan region (500 spp.), western North America (ca. 400–450 spp.) and along the Andes in South America (ca. 100 species) (Kazempour Osaloo *et al.*, 2003). *Astragalus* is diverse also in the Mediterranean climatic regions of southern Europe (including Asia Minor) and the New World's western coasts.

Although the diploid chromosome number is 2n=16 for the genus, high numbers of tetraploid, hexaploid and octoploid taxa were reported in cytological studies throughout the world in *Astragalus* (Ledingham, 1960; Ledingham and Rever, 1963; Masoud *et al.*, 2009). This suggests an important role of polyploidisation in the evolution of the genus. In addition to polyploidy, adaptive radiation on a global scale (Kazempour Osaloo *et al.*, 2003) probably made a significant contribution to the great number of species and high rates of endemism in *Astragalus*. A third factor that might have a significant role in species diversification is the rareness or even lack of hybridisation (Liston, 1992; Judd *et al.*, 2008; Kazemi *et al.*, 2009) in combination with widespread autogamy (D. Podlech, *personal communication* (2010)), which could have facilitated the isolation of species. As no evidence was reported yet for hybridisation in *Astragalus*, one could assume that neither hybridisation nor allopolyploid speciation had contributed to speciation in the genus. These phenomena had otherwise profoundly shaped the evolution of angiosperms (Grant, 1981; Soltis and Soltis, 2009).

Section *Dissitiflori* of the genus *Astragalus* with special emphasis on the European species

Section *Dissitiflori* is among the most species-rich (app. 160 spp.) sections of the genus *Astragalus* and is the largest section in the otherwise paraphyletic subgenus *Cercidothrix* which encompasses the bifurcate-haired Astragali (Ranjbar, 2004). This section was established by A. P. de Candolle in 1825. Later, Bunge (1868) introduced the same section under the name of *Xiphidium*, unaware that this had already been done by de Candolle (Ranjbar, 2004). Lectotypification was accomplished by Podlech (1990). As a result, the lectotype of this section is *Astragalus varius* S. G. Gmel.

Sect. *Dissitiflori* has its main centre of speciation in the mountains of Central Asia. Derived lineages of this section spread far to the north, east, west, and south (Goncharov *et al.*, 1946), extending the section's range to whole Eurasia (Ghahremani-Nejad, 2004). In Europe the section exhibit a disjunct distribution area and comprises some 50 species, while the Flora of Romania counts 7-8 taxa (Ciocârlan, 2009). Several narrow endemic *Astragalus* species of Europe are found in this section, e.g. *A. hispanicus* Coss. ex Bunge in Spain, *A. aquilanus* Anzal. and *A. vesicarius* L. subsp. *pastellianus* Pollini (Arcang.) in Italy, *A. peterfii* Jáv. in Romania, A. *tarchankuticus* Boriss. in Ukraine. Sect. *Dissitiflori* is considerably divers in the Balkans and Asia Minor.

According to Bunge (1868) species of this section are shrubby, sub-shrubby, or herbaceous perennials with well-developed stems coated with bifurcate or more rarely to subbasifixed hairs. Leaves are imparipinnate; stipules are free or adnate to petiole, otherwise free from each other. Racemes are mostly loose, borne on a well-developed peduncle. Bracteoles are absent. Flowers are shortly pedicellate. Calyx is tubular. Petals are glabrous; wing limbs are mostly rounded at apex to retuse or slightly emarginate. Legumes are sessile, or more rarely stipitate, 2-locular; valves are leathery. Photos of selected taxa (focusing on their inflorescences) of the European members of *Astragalus* sect. *Dissitiflori* are presented on Fig. 3.

Polyploidisation is thought to have a significant role in the diversification of species belonging to section *Dissitiflori* (Sytin, 2009).

Relationships (either phylogenetic or morphological) are scarcely known within the European taxa of sect. *Dissitiflori*, and the taxonomy of the section is blurred by the controversial nomenclatural history of several taxa. With the exception of the study of Zippel and Wilhalm (2009) – in which they examine the origin of different populations of a *Dissitiflori* taxon (see chapter 3.5) – there is no molecular phylogenetic study published addressing the question of species' relationships in this section.

Taxonomic uncertainties within the European taxa of *Astragalus* sect. *Dissitiflori*

One of the taxonomically problematic species groups in sect. *Dissitiflori* is the *A. vesicarius* L. aggregate which, according to its most recent circumscription including Europe but disregarding the former Soviet Union, includes three subspecies: *A. vesicarius* L. subsp. *vesicarius* (incl. *A. albidus* Waldst. & Kit.), *A. vesicarius* subsp. *pastellianus* and *A. vesicarius* subsp. *carniolicus* (A. Kern.) Chater (Podlech, 2008). *A. vesicarius* subsp. *vesicarius* has the widest but more disperse distribution, stretching from the Iberian Peninsula to Ukraine, whereas *A. vesicarius* subsp. *pastellianus* is endemic to north-western Italy, and *A. vesicarius* subsp. *carniolicus* is confined to north-western Italy, Croatia and Slovenia. Besides the above mentioned

three subspecies, two taxa were repeatedly affiliated with (or even subsumed within) the *vesicarius* group: *A. pseudoglaucus* Klokov and *A. tarchankitucus* Boriss. *A. pseudoglaucus* is confined to the North-western part of the Black Sea region (coastal regions of Bulgaria, Romania, Ukraine and probably Moldova), and *A. tarchankuticus*, grows exclusively in the Crimea, on the Tarhankut Peninsula, from where it was originally described (Borissova, 1951). In Flora Europaea, they are treated as 'probable' subspecies of *A. vesicarius* (Chater, 1968). Later, the new combination *A. vesicarius* subsp. *pseudoglaucus* (Klokov) Ciocârlan has been introduced (Ciocârlan and Sârbu, 2001). More recently, *A. tarchankuticus* was synonymised under *A. albicaulis* D.C. whereas *A. pseudoglaucus* was synonymised with *A. vesicarius* (Podlech, 2011), thus corroborating the view of Ciocârlan and Sârbu (2001) on this taxon based on morphological grounds.

The problems of origin and relationships of the Romanian narrow endemic *Astragalus peterfii*

An emblematic species of Romanian nature conservation, *Astragalus peterfii* Jáv. is an endemit of the Transylvanian Lowland (Câmpia Transilvaniei, Romania). It was discovered in 1916 by the bryologist Márton Péterfi near the village Suatu (County Cluj), and the species was dedicated to him by the author of the species, Sándor Jávorka. A botanical reserve was designated to preserve the species as early as in 1932 at the *locus classicus* by Professor Alexandru Borza. For several decades this single site remained the only known occurrence of Péterfi's Milk-vetch. An additional population was discovered in 1962, a few kilometres away near village Căianu (Bădărău *et al.*, 2000). *Astragalus peterfii* figures on several red lists and international conservation agreements (Habitats Directive, IUCN Red List, Bern Convention and Romanian National Red List). It is an octoploid (2n=64) species (Ledingham and Rever, 1963) of section *Dissitiflori*.

The problem of the origin and relatedness of *A. peterfii* has provoked a series of hypotheses, assessed mainly using morphological observations. The first remark on this subject was made by the author of the species itself. According to Jávorka (1916), *A. peterfii* is closely related to *A. glaucus* M. B., which grows near the city of Odessa (the Ukraine). Studying the genetic variability of the species using biochemical analysis, Borza (1998) found a pattern of isoenzymes suggesting an allopolyploid origin of *A. peterfii*. He also compared *A. peterfii* with 10 other Romanian *Astragalus* species based on protein polymorphism in order to assess phylogenetic relationships among them (Borza *et al.*, 1994). None of the investigated species turned out to be closely related to *A. peterfii*; an unsurprising result as none of the other investigated taxa belonged to section *Dissitiflori*. Pânzaru (2006) came forward with a new theory

concerning the taxonomy of the species claiming that *Astragalus peterfii* was a synonym of *Astragalus vesicarius* subsp. *pastellianus*, the endemic northern Italian subspecies.
The above short account confirms the interest in the problem of the origin and phylogenetic relationships of the Péterfi's Milk-vetch.

The surveyed taxonomic literature is suggestive for the presence of a putative species complex in *Astragalus* sect. *Dissitiflori* potentially comprising *A. pseudoglaucus*, *A. tarchankuticus*, *A. peterfii* and members of *A. vesicarius* s.l. Disentangling species complexes is often difficult and requires integration of information from different sources, e.g. plastid and nuclear genome, or morphometrics.

Astragalus molecular phylogenetic studies of the last two decades first were aimed at the circumscription of the whole genus, including the pioneer works which reconstructed main extrageneric affinities and intersectional relationships within it (Wojciechowski *et al.*, 1993; Wojciechowski *et al.*, 1999; Kazempour Osaloo *et al.*, 2003, 2005; Wojciechowski, 2005; Kazemi *et al.*, 2009; Zhang *et al.*, 2009). Nonetheless, after these earlier studies we are now poised to start to understand shallower levels of phylogenetic relationships, as exemplified by the study of Rihai *et al.* (2011) breaking new grounds in *Astragalus* in this respect. This tendency was maintained by Javanmardi *et al.* (2012) who reconstructed the phylogeny of *Astragalus* sect. *Alopecuroidei*. The present thesis also tries to continue the aforementioned trend by attempting to reconstruct the phylogeny of the European members of *Astragalus* sect. *Dissitiflori*.

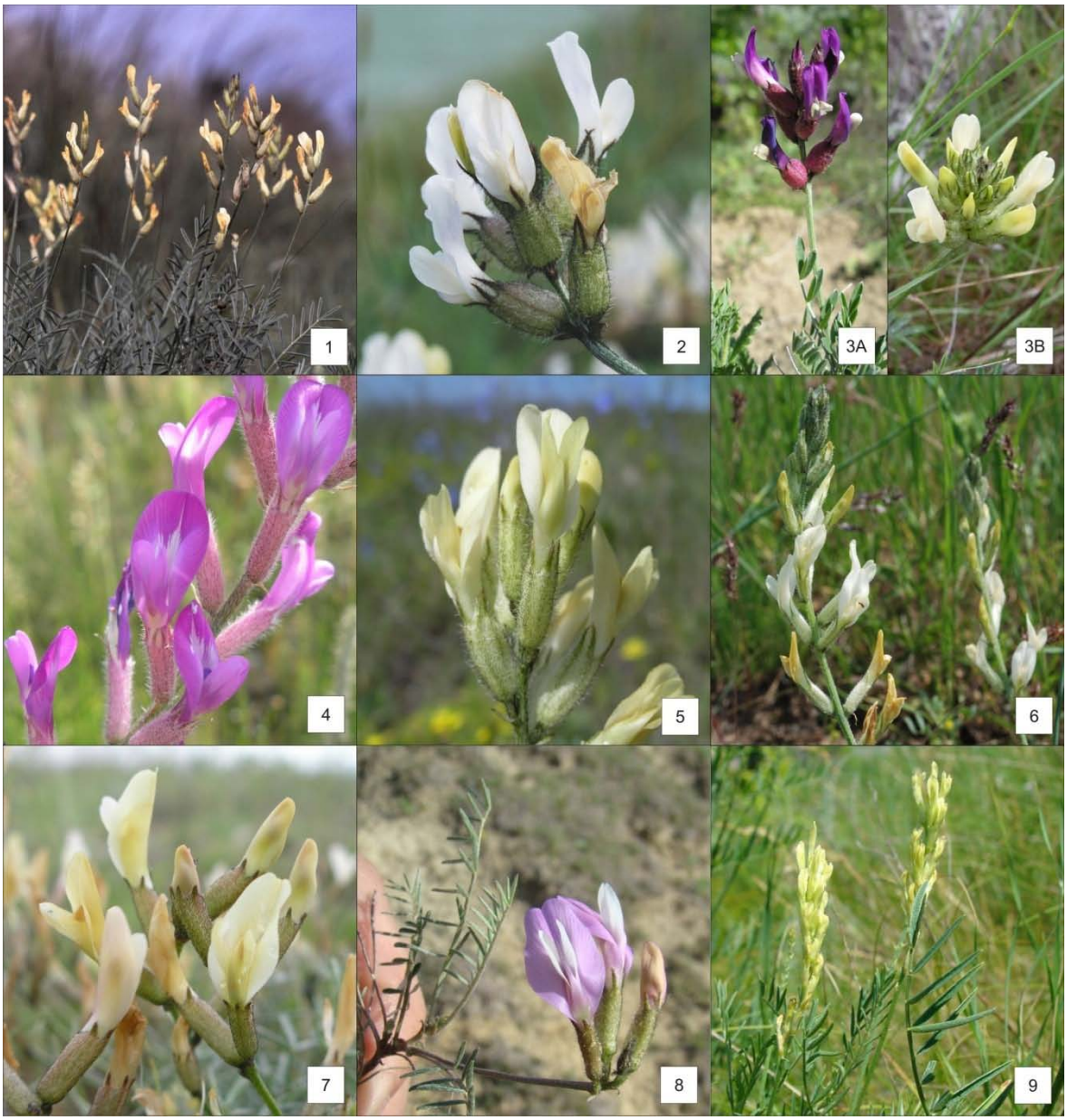

Fig. 3. Inflorescence images of selected taxa of the European members of *Astragalus* sect. *Dissitiflori*: 1 - *A. peterfii*, 2 - *A. pseudoglaucus*, 3A - *A. vesicarius* subsp. *albidus*, 3B - *A. vesicarius* subsp. *pastellianus*, 4 - *A. varius*, 5 - *A. tarchankuticus*, 6 - *A. pallescens*, 7 - *A. ucrainicus*, 8 - *A. subuliformis*, 9 - *A. asper*. (Photos by László Bartha except for the photo of *A. peterfii* by Attila Molnár V.)

Main objectives of the thesis

The principal goals of the thesis were as follows:

- to find and/or develop molecular phylogenetic markers for the genus *Astragalus* which are sufficiently variable to resolve relationships at the sectional level;
- to provide the first phylogenetic hypothesis for the European taxa of the genus *Astragalus* section *Dissitiflori* based on plastid sequences;
- to make an approach for estimating the section's timing of diversification in Europe;
- to test for the presence of reticulate speciation in the section/genus employing nuclear DNA markers;
- to explore the molecular phylogenetic relationships of the narrow endemic *Astragalus peterfii*;
- to try to disentangle a putative species complex with the aid of the plastid phylogeny, nuclear sequence data and multivariate morphometrics;

Materials and methods

Taxon sampling

The sampling design aimed at including as much European taxa of the section *Dissitiflori* as possible. For creating a better phylogenetic framework the sampling strategy included a limited number of taxa from Asia (mainly Asia Minor, the Caucasus and Middle Asia). For selection of taxa, the classifications of Podlech (2008, 2011) and Sytin (2009) were followed. With the exception of one relevant species – *A. apollineus* Boiss. from South Greece – I managed to collect material of almost all European representatives of the *Astragalus* section *Dissitiflori*. As for outgroup selection, I took into consideration the previously published plastid *ndh*F gene based *Astragalus* phylogenies of Kazempour Osaloo *et al.* (2003, 2005). I included three taxa from sections closely related to sect. *Dissitiflori*: two from sect. *Erioceras* (*A. reduncus* Pall. and *A. subarcuatus* Popov.) and one from sect. *Trachycercis* (*A. dolichophyllus* Pall.). To ensure the correct rooting of phylogenetic trees, a more distantly related outgroup was also selected. This was *A. glycyphyllos* L. (sect. *Glycyphyllus*) which was definitely outside of clade F (in the so called clade C sensu Kazempour Osaloo *et al.* (2003)). For the molecular clock analyses I included *Oxytropis pilosa* L. also due to the previously published (Wojciechowski, 2005) approximate age of the *Oxytropis–Astragalus* split which could be used as a prior setting in the analyses.

Plant material and DNA extraction

For genomic DNA extraction both herbarium and field collected material was used. In the latter case, leaves were dried and stored in silica gel until DNA extraction. In case of one sample of *Astragalus vesicarius* subsp. *pastellianus*, DNA was purchased from DNA bank (Botanical Garden, Berlin-Dahlem). ZR Plant/Seed DNA Kit (Zymo Research) was used for genomic DNA extraction.

Molecular marker development and marker selection

The selection of plastid markers to be used was preceded by a pilot screen of five candidate regions (*ndhF-rpl32*, *rpl*32-*trn*L, *trn*H-*psb*A, *acc*D-*psa*I and *pet*A-*psb*J) with the primers retrieved from Shaw *et al.* (2007). Amplification and sequencing of *trn*H-*psb*A, *acc*D-*psa*I and *pet*A-*psb*J has been successful with the universal primers retrieved from the literature, amplification of *ndhF-rpl32* and *rpl*32-*trn*L, however, has been poor or unsuccessful, but enough for the design of new, more efficient, *Astragalus* specific forward and reverse primers for these two regions.

Additionally, new (mainly Fabaceae specific) primers were designed for an approx. 1.5 kb region from the 3'-end of the plastid *ycf*1 region using as a guide the sequences of *Cicer* L.,

Glycine L., *Medicago* L. and *Phaseolus* L. *ycf*1 retrieved from GenBank. After the first pilot amplifications and sequencing, a more reliable *Astragalus* specific reverse primer was designed and used for all subsequent reactions.

Among the six plastid regions tested, *ycf*1, *ndh*F-*rpl*32 and *rpl*32-*trn*L seemed to be the most variable (data not shown); therefore these markers were chosen for the subsequent phylogenetic work.

As for the nuclear markers, I started to work with the popular nrITS region, and also attempted to develop a single-copy nuclear gene marker for *Astragalus*. This was the LEAFY (LFY) which had been found to be single-copy in several other plant groups (Zimmer and Wen, 2012), and references therein). I used the primers ITS5 (forward) and ITS4 (reverse) of White *et al.* (1990) for amplification of nrITS region, while for the second intron or first intron + second exon + second intron region of LEAFY new primers were designed on the basis of Fabaceae sequences retrieved from *GenBank*.

The single copy feature of LEAFY was tested in a few diploid taxa (e.g. *A. hispanicus*, *A. varius*, *A. vesicarius* subsp. *albidus*). After the first pilot PCR amplifications, sequencing and cloning, it turned out that this region was not single copy in *Astragalus*, therefore, its use was omitted from subsequent work. Three copies/paralogs of this gene have been retrieved from selected diploid *Astragali* and their identity was confirmed by BLAST search of the GenBank nucleotide database.

Plastid DNA markers were planned to be analysed usually from one specimen of species, while the nrITS region was attempted to be sampled for at least two specimens per species. Plastid DNA data was planned to be supplemented with nuclear DNA data for selected taxa of mostly uncertain taxonomy or unknown phylogenetic position (see Theoretical backgrounds section).

Cloning and sequencing

The direct sequencing of nrITS showed unambiguous double peaks in the majority of accessions suggesting the presence of more than one nrITS sequence type in the sample. In order to detect this intra-individual variability, I performed cloning using pGEM-T Vector System I (Promega) for ligation and GeneJET Plasmid Miniprep Kit (Fermentas) for plasmid isolation. nrITS was cloned from 23 accessions (belonging to ten taxa) originating from 21 populations with an average of 7.8±2.5 cloned sequences per specimens.

All DNA sequencing was performed at commercial service at Macrogen Inc. (South Korea).

Phylogenetic analyses

DNA sequence aligning

Sequences of *ndh*F-*rpl*32, *rpl*32-*trn*L and nrITS clones were directly exported from the chromatograms to FASTA format files using ChromasLite v.2.01 (Technelysium Pty). *ycf*1 fragments of the same taxonomic sample (obtained with different internal sequencing primers) were assembled using BioEdit (Hall, 1999).

Sequences were aligned in MEGA5 (Tamura *et al.*, 2011) by algorithm ClustalW followed by subsequent manual editions in case of the plastid DNA regions. In case of *ycf*1, during the manual correction of the alignment, codon positions were taken into consideration. Translation for this region was achieved in MEGA5 using the plant plastid code and resulted in an alignment containing no stop codons.

As suggested by Vignal *et al.* (2002) only those singleton mutations above 1% were considered to be single nucleotide polymorphisms (SNPs) in the nrITS matrix, others below this threshold limit were corrected manually. nrITS clone sequences with uncommonly long deletions and/or many point mutations were regarded as pseudogenes and were excluded from subsequent work.

Phylogenetic tree reconstructions based on plastid sequences

To keep matters simple, separate plastid data matrices have been analysed only under parsimony criteria in order to get insight into the resolving power of these markers and reveal putative conflicting signals between them. Therefore, topologies of the strict consensus trees derived from the MP analysis of the separate plastid data matrices were at first compared. My strategy was to apply different tree reconstruction methods (maximum parsimony (MP), Bayesian inference (BI) and maximum likelihood (ML)) only to a combined final dataset which lacks taxa with strongly contrasting positions between the strict consensus trees of the separate plastid regions (either statistically supported or not). Moreover, such a combined final dataset should easily pass (p value > 0.05) the incongruence length difference ILD test (Farris *et al.*, 1994) which is widely used to uncover discrepancies between different DNA regions which are going to be concatenated.

The ILD test (as implemented in *PAUP** 4.0b10 (Swofford, 2002)) was employed using 100 replicates of heuristic searches with TBR branch-swapping.

MrModeltest v2. (Nylander, 2004) was used to select the nucleotide substitution models for the three plastid DNA regions using the AIC criterion. Bayesian analysis was carried out on a partitioned dataset with the nucleotide substitution models listed in Table 5, involving two

simultaneous runs of 10.000.000 generations of Monte Carlo Markov chains by saving every 10.000 tree into a file. Each run employed four simultaneous chains (one cold and three heated with a heating parameter of 0.2). The 50% Majority Rule consensus phylogram was generated in MrBayes using 25% as "*burn-in*".

Maximum parsimony analysis was run in *PAUP** and relied on heuristic search using 1000 random addition of sequence replicates and tree bisection-reconnection (TBR) branch swapping with MULTREES option in effect, MAXTREES set to 50.000 (without possibility of 'promt' and 'auto-increase') and a limit of ten trees retained for each replicate. Characters were weighted equally and gaps were treated as missing data. The statistical robustness of tree branches was estimated via bootstrapping method; 1000 pseudo-replicates were performed in *PAUP** with MAXTREES (re)set to 1000 and with the retention of one tree per replicate.

Maximum likelihood analysis relied on RAxML (Stamatakis, 2006) using the RAxML GUI version 1.2 (Silvestro and Michalak, 2011) under the GTR + Γ model of sequence evolution. Nodal support values for the ML topology were estimated using the rapid bootstrap algorithm implemented in RAXML employing 100 replicates (Stamatakis *et al.*, 2008).

All heuristic searches and Bayesian analyses were run at the University of Oslo Bioportal (https://www.bioportal.uio.no/). Phylogenetic trees were visualised and edited in TrieeView (Page, 1996), FigTree version 1.3.1 (A. Rambaut; http://tree.bio.ed.ac.uk/software/figtree/), MrEnt (Zuccon and Zuccon, 2006) and CorelDRAW X3.

Network analyses of nrITS clone sequences

Because of the large number of clones and presence of diverged nrITS paralogs in some of the samples, phylogenetic network approaches rather than hierarchical tree-based ones were at first used in the nrITS analyses.

The program Collapse v.1.2 (Posada, 2006) was used for defining unique sequence types (ribotypes) in the nrITS dataset and assess the distribution of these ribotypes within and between accessions. The defined unique ribotypes were then included in parsimony network analysis

A distance network analysis (split graphs) was carried out on the ribotype matrix in order to deliniate putative groupings within the nrITS dataset and to assess confidence to groupings using Bootstrapping with 1000 replicates. The phylogenetic network constructed for this purpose was based on NeighborNet (NN) algorithm (Bryant and Moulton, 2004) as implemented in SplitsTree v.4.10 (Huson and Bryant, 2006) using uncorrected pairwise (p) distances, excluding both constant and non-informative characters.

Results

Sequencing success and sequence alignments

Plastid *ycf*1, *ndh*F-*rpl*32 and *rpl*32-*trn*L

Amplification and sequencing of *ycf*1 has been successful in case of each (47) targeted *Astragalus* taxa plus *Oxytropis pilosa*. The partial *ycf*1 region flanked by the amplification primers has been completely recovered by assembling the fragments obtained with the internal sequencing primers and belonging to the same taxonomic sample. The length of the assembled region varied from 1460 base pairs (bp) (*A. argyroides*) to 1517 bp (*A. zingeri*).

Amplification of *ndh*F-*rpl*32 and *rpl*32-*trn*L failed in case of several, mainly herbarium samples leading to a maximum number of 11 missing sequences in case of *rpl*32-*trn*L. Sequence characteristics of the plastid regions and assessments of MP heuristic searches are summarised in Table 5.

nrITS

The original aligned nrITS matrix contained 181 sequences, 598 characters and 56 variable sites. Only one base long, putative parsimony-informative indels were found at three positions. The program Collapse retrieved 54 unique ribotypes from the original 181 sequences.

Table 5. Sequence characteristics and assessment of maximum parsimony heuristic searches of plastid datasets

	ycf1	ndhF-rpl32	rpl32-trnL	combined
sequence lengths	1460–1517	630–704	738–809	–
alignment length	1553	803	915	3271
no. of missing sequences	0	7	11	18
no. of variable sites	128	51	45	159
no. of parsimony inf. sites	45	11	9	65
equally most parsimonious trees	–	–	–	27559
tree length	–	–	–	276
CI	–	–	–	0.8696
RI	–	–	–	0.9130
evolutionary model selected (under AIC)	GTR+G+I	GTR	GTR	GTR+G+I/GTR/GTR

Phylogenetic tree analyses of plastid sequences

Heuristic search of the combined, complete plastid dataset under parsimony criteria resulted in 27559 equally most parsimonious trees with 276 steps in length. The A and B clades (retrieved already by previous analyses of the separate data matrices) have been recovered with high Bootstrap support percentages (87% and 96%, respectively). The inclusion/exclusion of *A. medius*, *A. reduncus* and *A. ucrainicus* into clade A was not consistent throughout the parsimony analyses of the separate plastid datasets. On this tree, however, these species are firmly excluded from the highly supported clade A. The A and B clades together with a third clade (comprising *A. fissuralis*, *A. subuliformis* and *A. ucrainicus*) plus three additional species (*A. balkaricus*, *A. medius* and *A. reduncus*) form a rather unsupported basal polytomy. This polytomy is sister to *A. argyroides* which itself branches after *A. dolichophyllus* implying that *A. argyroides* constitutes the first diverging lineage of the *Dissitiflori* taxa. Clade A breaks up into several subclades a part of which was tentatively designed in a successive (i.e. clades A, A1, A1.1) manner. Clade B was less resolved than clade A.

There are both within clades A and B several surprising relationships which worth noting. *Astragalus subarcuatus* (traditionally classified under sect. *Erioceras*) was inferred with high statistical certainty (BS=94%) as sister to *A. temirensis*. Another *Erioceras* taxon (*A. reduncus*) is part of the already mentioned basal polytomy.

The most surprising relationship from within clade B consists of the highly supported (BS=91) sister relationship of two taxa.

Careful examination of the taxa content of clades (and the species distribution) led to the recognition that clades A, B and B1 show broad geographical structure: clade A has mostly East European distribution ('eastern clade') while clade B (excluding clade B1) has mostly Mediterranean distribution ('south-western clade'), whereas clade B1 is definitely confined to East Europe.

The topology of consensus phylogram derived from the Bayesian analysis is broadly congruent with that of the MP strict consensus tree. Clades A, B and B1 are highly consistent between the parsimony and Bayesian trees as far as the taxa content and statistical support of internal relationships are considered. The long branch of the Bayesian tree leading to the polytomy of clade B is conspicuous as compared with the rest of the internal branches of the tree.

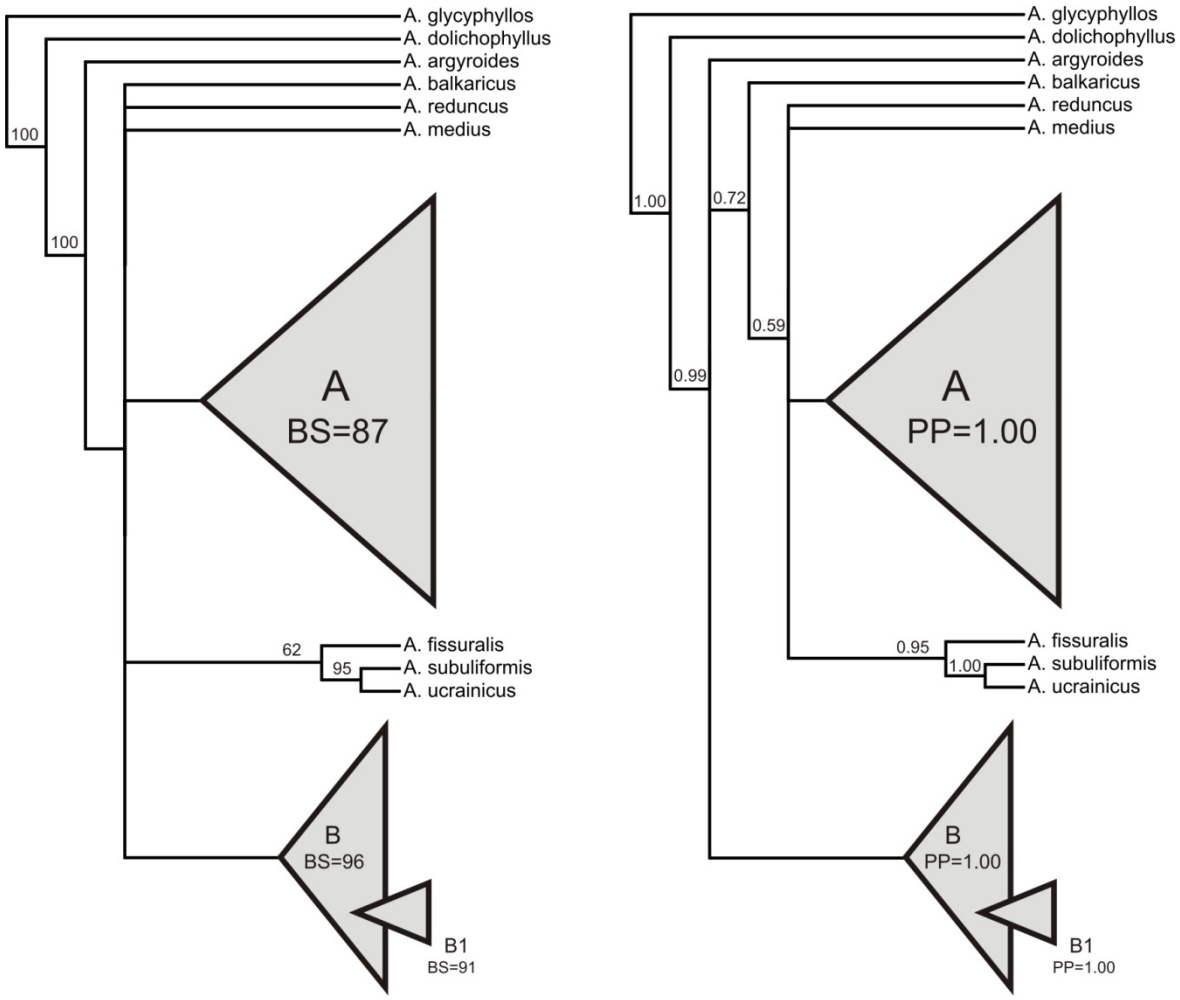

Fig. 9. Schematic comparison of the maximum parsimony and Bayesian consensus topologies. Numbers adjacent to branches/clades represent MP Bootstrap support percentages and Bayesian posterior probability values, respectively. Taxa content and internal relationships of clades A, B and B1 are confidential but are presented in the thesis.

Network analysis of nrITS sequences

The NeighborNet analysis of ribotypes (Fig. 13) indicated a non-tree-like evolution of the sequences (with fit value 97.74 of the resulting split graph) hinting at more complex evolutionary processes acting in the background. Seven groups could be separated on the network (labelled from A to G), which corresponded with the groupings delimited on the TCS network. Therefore, sequence content of the groupings corresponded between the TCS and NeighborNet networks.

Bootstrap support values were high (BS=89-100%) for five out of the seven groups and low in case of ribotype groups E (BS=64%) and F (BS=66%).

Group A was topologically the most segregated group on the NeighborNet split graphs and statistically was the best supported (BS=100%) among all groupings. This is concordant with its outlier position from within the TCS network.

nrITS sequences retrieved from *A. asper*, *A. ucrainicus* and *A. vesicarius* each belonged to one single ribotype group (C, G, and E, respectively), while sequences of the rest of the five species clustered into 2–4 different groups. *Astragalus peterfii* has the single highest number of ribotype groups identified within it; nonetheless, incomparably high number of clones was sequenced in this species.

The number of different species which shared the same ribotype group varied from two to four. Sequences from ribotype groups C and F belonged exclusively to one species (*A. asper*, and *A. varius*, respectively) while the rest of the ribotype groups were shared by different species. Interestingly, groups A and B are the dominant ribotype groups of *A. pallescens*, *A. peterfii* and *A. pseudoglaucus*.

The frequency of nrITS sequences belonging to a given ribotype group within the species has shown considerable differences (Table 9). Some groups were more characteristic for one species than another by including more sequences from the given species, e.g. ribotype group E encompassed the sequences retrieved mainly from *A. vesicarius* (i.e. 'vesicarius-like' ribotype group).

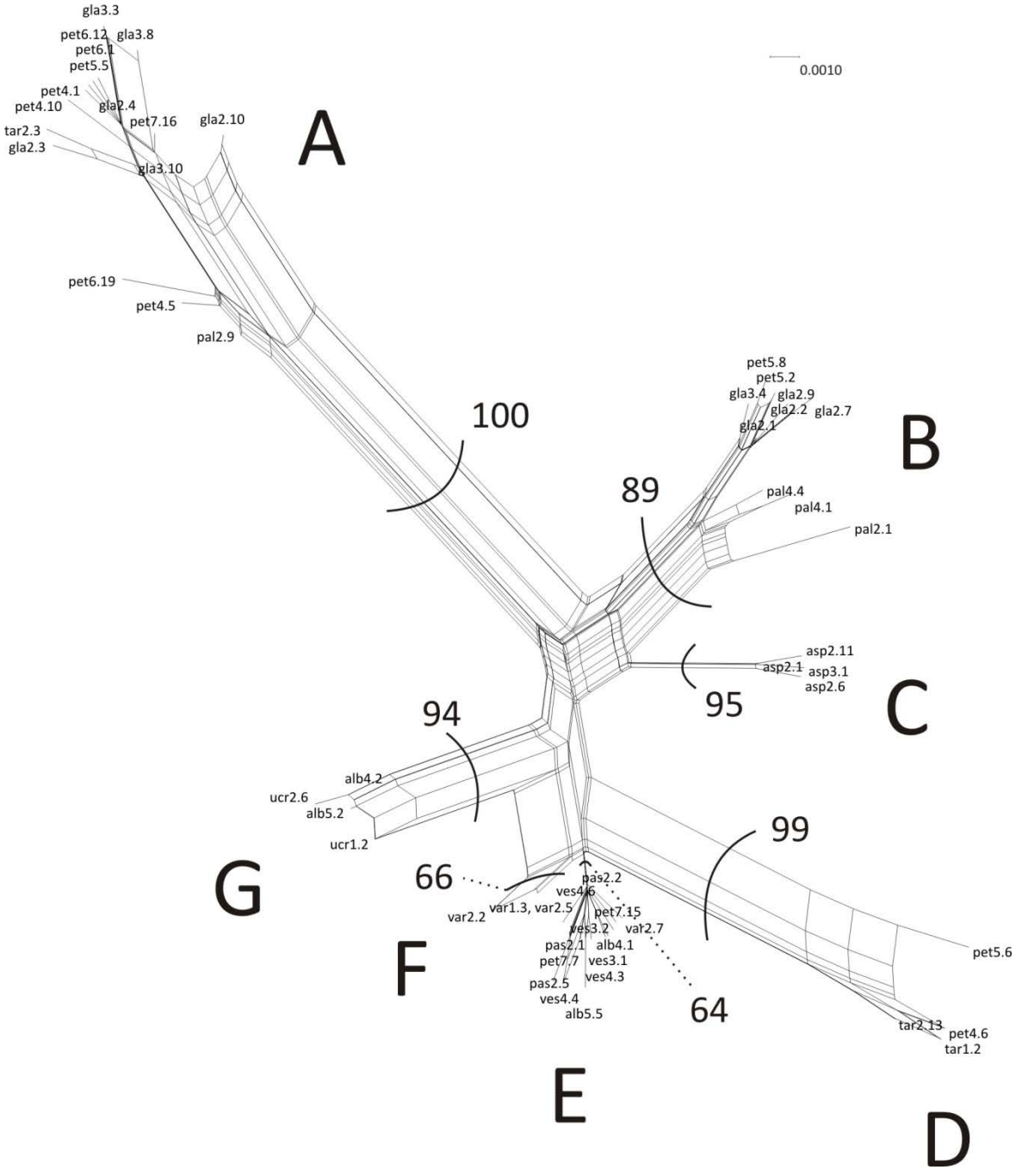

Fig. 13. Phylogenetic network based on NeighborNet analysis of the ribotypes. Sequences are grouped together to genetically similar groups as represented by the capital letters A-G. Values represent Bootstrap support percentages of 1000 Bootstrap replicates.

Discussion

Phylogeny of *Astragalus* section *Dissitiflori* in Europe: implications to biogeography and timing of diversification

The first plastid based phylogenetic hypothesis for the European taxa of the genus *Astragalus* sect. *Dissitiflori* relied on a comprehensive taxon sampling and shed lights to several peculiar aspects of this groups' diversification. The broad geographic structure revealed for clades A, B and B1 unravelled three main biogeographic patterns for the section: a major East European and Asian distribution represented by clade A, a mostly Mediterranean distribution (but extending to Central-West Europe and the Caucasus) represented by clade B (excluding clade B1) and a minor branching of clade B (i.e. clade B1) representing the East European and Asian distribution of two sister taxa. At first it was tempting to hypothesise a more ancient diversification for clade A as compared with clade B, and a more rapid diversification for clade B as compared with clade A. Facts arguing for these predictions were as follows: the greater resolution (with well supported internal relationships) of clade A as compared with the unresolved polytomy of clade B and the long branch leading to clade B as compared with the sort branches distributed from the basal *Dissitiflori* split towards the tips of clade A (reflected by e.g. the ML phylogram of the plastid dataset).

Polytomies (lack of resolution) of phylogenetic trees (especially at deep nodes) have usually been regarded as consequences of rapid speciation, use of inadequate (i.e. insufficiently variable) markers or character conflicts resulting from multifurcate speciation (Riggins and Seigler, 2012), and references therein). In case of clade B, multifurcate speciation can be excluded because plastid markers belong to the same genetic linkage group and are likely to have the same phylogenetic history. The combination of three plastid markers used has been effective in one (A) clade but unsuitable in another (B) clade of a moderately large taxonomic group. This would imply a more rapid speciation of taxa from clade B as compared with species from clade A or lineage (clade) specific heterogeneity in the rate of nucleotide substitutions, a phenomenon not inexistent in plants (Weng *et al.*, 2012), and references therein). I attribute low sequence divergence (variation) to the generally less unresolved feature of clade B, a fact confirmed also by pairwise distance analysis of sequences. Pairwise distances – as calculated in Mega 5.0 under the Maximum Composite Likelihood model – ranged between 0.001–0.005 (overall mean 0.003) in case of clade B, while in case of clade A this value varied between 0.001–0.012 (overall mean 0.005).

Finally, a shorter time frame in which clade B could have diverged was not confirmed by the molecular clock analysis. According to one of the resulted chronograms diversification of both

clades underwent during the Pleistocene. Within clade B the youngest age (crown) had been assigned to clade B1 (around 500.000 years) which is in concordance with a hypothesised 'secondary' ramification of clade B accounting for the diversification of two taxa in East Europe/Asia.

Non-monophyly of sections *Dissitiflori* and *Erioceras*

Both species of sect. *Erioceras,* initially selected as outgroups, are nested in sect. *Dissitiflori* making the latter non-monophyletic. *Astragalus subarcuatus* is deeply nested within clade A and was resolved with high statistical certainty (MP BS=94%, BI PP=1.00, ML BS=98%) as sister to *A. temirensis*, whereas *A. reduncus* is among the earliest deriving species of plastid phylogenies but branching after *Dissitiflori* taxa (e.g. *A. argyroides* and *A. balkaricus*) on the basis of the maximum likelihood phylogram. In order to acquire monophyly for sect. *Dissitiflori*, both taxa should be transferred into this section. In fact, more thorough investigations are needed on the delimitation of these sections prior to their redefinition. More taxa should be included from section *Dissitiflori* from its center of diversification and sect. *Erioceras* should also be more comprehensively sampled.

In case of an eventual redefinition of sect. *Dissitiflori*, consideration should be taken on different alternatives leading to a more broadly vs. more narrowly defined section. A more broadly treatment would imply the inclusion of *Erioceras* taxa, whereas splitting of the currently recognised section into sections/subsections would imply taxonomic assignation e.g. to the very compact clade B. Following monophyly, 'morphological diagnosability is an important subsidiary criterion for the classification of groups' (Chatrou *et al.*, 2012), thus until a morphological character or a combination of characters supporting this clade is not identified, any attempts for assigning a taxonomic rank to clade B seems meaningless. On the other hand, taxonomic decisions should not be based exclusively on plastid DNA phylogenies. Creating a backbone phylogeny based on at least one low-copy nuclear marker it is highly recommended.

Mapping different states of a set of morphological characters on clade B do not resulted in the recognition of at least one constant character which could potentially serve as a synapomorphy for this clade when comparing its members e.g. with taxa of clade A. As a next step, such additional (qualitative) morpho-anatomical characters should be screened which are thought to be not influenced by environmental factors. Pollen micromorphology meets this criterion and therefore it would be worth surveying in taxa of clade A vs. clade B.

Fig. 17. Photos of leaflets of selected taxa of sect. *Erioceras* (1 - *A. reduncus*; 2 - *A. subarcuatus*) and sect. *Dissitiflori* (3 - *A. temirensis*; 4 - *A. pseudoglaucus*; 5 - *A. vesicarius*; and *A. peterfii*) highlighting the difference in the pubescence of vegetative characters between the two sections.

Reticulate speciation and incomplete concerted evolution of nrITS in *Astragalus* section *Dissitiflori*

Although some major contributions to the phylogeny of *Astragalus* and its sister genus *Oxytropis* are based partially or totally on nrITS (Wojciechowski *et al.*, 1999; Jorgensen *et al.*, 2003; Kazempour Osaloo *et al.*, 2003, 2005; Wojciechowski, 2005; Kazemi *et al.*, 2009; Archambault and Strömvik, 2012), these studies do not report the cloning of this marker. Thus, the present study can be considered to be the first work reporting serious paralogy of nrITS in a group of *Astragalus* by utilising deep-cloning. Scherson *et al.* (Scherson *et al.*, 2005) screened novel nuclear loci for reconstructing phylogenies at low taxonomic levels in New World *Astragalus*. They confirmed by cloning the presence of different copies of two nuclear loci (ARG10 and FENR) and SNPs in the nuclear locus tRALS in some taxa of New World *Astragalus*. This

pattern, however, was interpreted as consequence of duplication events, and presence of alleles at the given loci without phylogenetic significance. Therefore, my interpretation for the presence of paralogy in the nrITS in *Astragalus* is the first taking reticulation as the possible source of paralogy into consideration. Moreover, the reticulate structure of nrITS in the polyploid *A. pallescens* (2n=32 (Philippov *et al.*, 2008)), *A. peterfii* (2n=64 (Ledingham and Rever, 1963)) and *A. pseudoglaucus* (2n=64 (Pavlova and Kozhuharov (1993), under *A. glaucus*)) is suggestive for their allopolyploid origin, thus allopolyploidy in *Astragalus* is the first time hypothesised here based on molecular data.

The reticulate structure of the nrITS data set strongly suggests that evolutionary processes different from dichotomous splitting of lineages (e.g. merging) took place in section *Dissitiflori*. The parallel persistence of the dominant ribotype groups A and B in a single genome suggests a retarded concerted evolution (Campbell *et al.*, 1997) of nrITS. The A and B copies might persist in the putative parental progenitors or in their descendants forming a unique ribotype group (a single group per species, according to the present concept). Retardation or incompletion of concerted evolution has long been known in other plant groups (both within diploids and polyploids). Classical examples for incomplete concerted evolution with respect to the nuclear ribosomal DNA includes *Amelanchier* (Campbell *et al.*, 1997), *Arabidopsis suecica* (O'Kane *et al.*, 1996), *Brassica napus* (Bennett and Smith, 1991), *Paeonia* (Sang *et al.*, 1995) but new examples are continuously being discovered and it seems that incomplete homogenization of nrITS is the rule rather than the exception (Liu *et al.*, 2006). Factors such as the presence of different nrITS arrays at different chromosomes (e.g. due to allopolyploidy), asexual reproduction, and perennial habit (Sang *et al.*, 1995; Campbell *et al.*, 1997) may promote the maintenance of nrITS polymorphism (i.e. mitigation of unequal crossing over and gene conversion to complete concerted evolution (Hillis *et al.*, 1991)).

Origin and relationships of *A. peterfii*

My results suggest an allopolyploid origin of *Astragalus peterfii* due to the presence of four distinct nrITS ribotype groups in the genome of this octoploid species which is hardly compatible with an exclusive autopolyploid origin. As for the molecular phylogenetic relationships, the plastid phylogeny placed *A. peterfii* with *A. pallescens* and *A. tarchankuticus.* Out of these two species, however, only *A. pallescens* shares the dominant nrITS ribotype groups (type A and B) with *A. peterfii*, and moreover, in a seemingly similar ratio. Multiple Discriminant Analysis of 14 quantitative morphological characters from field-collected plants, however, revealed a substantial morphological overlap between *A. peterfii* and *A. pallescens*, and a complete distinctness between *A. peterfii* vs. *A. pseudoglaucus*, and *A. peterfii* vs. *A. tarchankuticus*. Considering the

above three types of information, it can be concluded that out of the studied taxon-set *A. pallescens* is the closest relative of *A. peterfii*. Creating a more powerfull hypothesis for the phylogenetic relationships of *A. peterfii* might require the use of low-copy nuclear genes, which are at least less susceptible to concerted evolution (Zimmer *et al.*, 1980; Hillis *et al.*, 1991).

Taxonomic implications to selected taxa of sect. *Dissitiflori*

The distribution of specific ribotype-groups within the species, in connection with the plastid phylogeny and multivariate morphometric data, allows drawing some conclusions on the taxonomy of those species for which taxonomic uncertainties have been highlighted in the Introduction chapter. For instance, all nrITS sequences of *A. vesicarius* belong to ribotype group E, nrITS sequences of *A. pseudoglaucus* and *A. pallescens* are found mostly within groups A and B, but the latter two taxa differ in their placement on the plastid phylogeny and are segregated by multivariate morphometric analysis.

The synonymisation of *A. peterfii* under *A. vesicarius* subsp. *pastellianus* is supported neither with plastid, nor with nrITS sequence data. This, however, does not refute the fact that *A. vesicarius* might have contributed to the formation of my species of interest since ribotype group E (in a small portion) was recovered also from *A. peterfii*.

The situation is somehow similar in the case of *A. pseudoglaucus* and *A. tarchankuticus*. *A. pseudoglaucus* was synonymised with *A. vesicarius* by Ciocârlan and Sârbu (2001) and Podlech (2011). In my opinion, the submerge of *A. pseudoglaucus* into *A. vesicarius* is misleading and is refuted by nrITS sequence data, though *A. vesicarius* might be involved in the formation of the species as plastid DNA donor. The striking difference between *A. pseudolaucus* and *A. vesicarius* concerning their nrITS structure and phylogenetic position is counterbalanced by their overlap in morphology which argues for cryptic genetic diversity in this taxonomic group.

The taxonomic independence of *A. tarchankuticus* is also supported with respect to *A. albicualis* with which it has been synonymised (Podlech, 2011). Both *A. tarchankuticus* accessions each have one nrITS clone sequence in ribotype group E and have most of their nrITS sequences in group G. Contrary to this, nrITS sequences of the two *A. tarchankuticus* samples are clustered mostly in groups A and D in a similar ratio.

Several factors could have led to the uncertain taxonomy of these species. Among the already revealed reticulation, parallelism in morphology might have also hampered the identification of hidden taxonomic richness in this group.

Another consequence of ribotype group differentiation is the molecular confirmation of a species complex within section *Dissitiflori*, formed by *A. pallescens*, *A. peterfii*, *A. pseudoglaucus,*

and *A. tarchankuticus*. All of these species have the ribotype group A as a presumed 'core' of their nrITS array and are known polyploids with the exception of *A. tarchankuticus*.

Phylogenetic utility of the combined use of *ycf*1, *ndh*F-*rpl*32 and *rpl*32-*trn*L regions in *Astragalus* phylogenetics

In the present study the plastid regions *ycf*1, *ndh*F-*rpl*32 and *rpl*32-*trn*L have been used in *Astragalus* phylogenetics for the first time. The primary goal in utilising these markers was to infer relationships at exclusively low taxonomic levels (i.e. at the intrasectional level). Therefore, their variability is hardly comparable with the already used plastid markers utilised at higher taxonomic levels (e.g. at generic or inter-sectional levels) in *Astragalus* phylogenetics: *trn*L intron (Wojciechowski *et al.*, 1999), *ndh*F gene (Kazempour Osaloo *et al.*, 2003), *trn*L-F region (Kazemi *et al.*, 2009). The recent comparison of Dong *et al.* (2012) involving 23 plastid regions (among others, *ycf*1 and *rpl*32-*trn*L, but omitting *ndh*F-*rpl*32) found *ycf*1 the most variable followed by *trn*K and *rpl*32-*trn*L, thus confirming the perspectives of these regions in plant phylogenetic inference.

In spite of the fact that *ycf*1+*ndh*F-*rpl*32+*rpl*32-*trn*L could not discriminate well diverged species like *A. tarchankuticus* from *A. peterfii*, it was sufficiently variable to delimitate several well supported clades within section *Dissitiflori* (see Results chapter) and even within the recognised species complex. In my opinion, the three plastid regions used here might have potential for further phylogenetic studies in *Astragalus,* e.g. for sectional delimitations.

General conclusions

- A combined phylogenetic analysis with three plastid regions sampled in the European taxa of the genus *Astragalus* sect. *Dissitiflori* provided sufficient resolution to draw definite conclusions on the group's diversification.

- The main feature characterising the group's plastid phylogeny is a deep evolutionary split leading to two main geographically structured clades: the 'eastern' clade encompasses species mainly of East European and Asian distribution; and the 'south-western' clade comprises species mainly from the Mediterranean (incl. Asia Minor), Central and West Europe and the Caucasus. The 'south-western' clade has a younger ramification which reflects the surprising close relationship between two taxa, both with contrasting East European and Asian distribution.

- Incongruence between the plastid phylogeny and the phylogenetic position of *A. vesicarius* s.l. on the nrITS phylogenetic tree strongly suggests that the 'eastern' and 'south-western' clades interacted as a result of past hybridisation.

- Section *Dissitiflori* is not monophyletic in its current definition because the two taxa included from sect. *Erioceras* are nested within it.

- The strongly reticulate structure of nrITS in a group of mostly polyploid taxa of controversial nomenclatural history (*A. peterfii, A. pallescens, A. pseudoglaucus, A. tarchankuticus*) argue for their allopolyploid origin. Reticulate structure of nrITS has been firstly revealed for the genus in the framework of this PhD thesis, and paralogy of a nuclear marker in *Astragalus* was explained with past hybridisation (i.e. reticulate speciation) for the first time here.

- Ribotype group differentiation, in combination with the plastid phylogeny and multivariate morphometric data, proved to be a useful tool in suggesting the taxonomic independence of several species of taxonomic instability in *Astragalus* sect. *Dissitiflori*.

Selected literature

Alfaro, M. E., and Holder, M. T. 2006. The posterior and the prior in Bayesian phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*. **37**: 19-42.

Álvarez, I., and Wendel, J. F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*. **29**(3): 417-434.

Archambault, A., and Strömvik, M. V. 2012. Evolutionary relationships in *Oxytropis* species, as estimated from the nuclear ribosomal internal transcribed spacer (ITS) sequences point to multiple expansions into the Arctic. *Botany*.

Bădărău, A. S., Dezsi, Ş., and Comes, O. 2000. Cercetări biogeografice asupra speciilor stepice-silvostepice de *Astragalus* L. din Depresiunea Transilvaniei (I) [A biogeographical study upon the steppe and steppe-woodland species of *Astragalus* L. from the Transylvanian Depression (I)]. *Studia Universitatis Babeş-Bolyai Series Geografica*. **45**: 117-130.

Baldwin, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: An example from the compositae. *Molecular Phylogenetics and Evolution*. **1**(1): 3-16.

Bennett, K. D., Tzedakis, P. C., and Willis, K. J. 1991. Quaternary refugia of north European trees. *Journal of Biogeography*. **18**: 103-115.

Bennett, R. I., and Smith, A. G. 1991. Use of a genomic clone for ribosomal RNA from *Brassica oleracea* in RFLP analysis of *Brassica* species. *Plant Molecular Biology*. **16**(4): 685-688.

Borissova, A. 1951. Astragali et Tragachantae Tauriae. *Botanicheskie Materialy Gerbariya Botanicheskogo Instituti Imeni V L Komarova Akademii Nauk SSSR*. **14**: 213-228.

Borza, T. 1998. Genetic polymorphism of some rare or endemic plants., Universitatea Babeş-Bolyai, Cluj Napoca, Romania.

Borza, T., Dragoş, N., and Clejan, A. 1994. Considerations on some species of *Astragalus* L. genus as revealed by electrophoretic study of seed storage proteins. *Contribuţii Botanice*. **97**: 102.

Bryant, D., and Moulton, V. 2004. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*. **21**(2): 255-265.

Bunge, A. 1868. Generis Astragali species Gerontogeae. Pars prior. Claves diagnosticae. *Mémoires de l'Académie impériale des sciences de St Petersbourg*. **11**(16): 1-140.

Calonje, M., Martín-Bravo, S., Dobeš, C., Gong, W., Jordon-Thaden, I., Kiefer, C., Kiefer, M., Paule, J. R. S., and Koch, M. A. 2008. Non-coding nuclear DNA markers in phylogenetic reconstruction. *Plant Systematics and Evolution*. **282**(3-4): 257-280.

Campbell, C. S., Wojciechowski, M. F., Baldwin, B. G., Alice, L. A., and Donoghue, M. J. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). *Molecular Biology and Evolution*. **14**(1): 81-90.

Chase, M. W., De Bruijn, A. Y., Cox, A. V., Reeves, G., Rudall, P. J., Johnson, M. A. T., and Eguiarte, L. E. 2000. Phylogenetics of Asphodelaceae (Asparagales): an analysis of plastid *rbcL* and *trnL-F* DNA sequences. *Annals of Botany*. **86**(5): 935-951.

Chase, M. W., Williams, N. H., de Faria, A. D., Neubig, K. M., Amaral, M. d. C. E., and Whitten, W. M. 2009. Floral convergence in Oncidiinae (Cymbidieae; Orchidaceae): an expanded concept of *Gomesa* and a new genus *Nohawilliamsia*. *Annals of Botany*. **104**(3): 387-402.

Chater, A. O. 1968. *Astragalus* L. *In* Flora Europaea. *Edited by* Tutin, T. G., Heywood, V. H., Burges, N. A., Valentine, D. H., Walters, S. M., and Webb, D. A. Cambridge University Press, Cambridge. pp. 108-124.

Chatrou, L. W., Pirie, M. D., Erkens, R. H. J., Couvreur, T. L. P., Neubig, K. M., Abbott, J. R., Mols, J. B., Maas, J. W., Saunders, R. M. K., and Chase, M. W. 2012. A new subfamilial and tribal classification of the pantropical flowering plant family Annonaceae informed by molecular phylogenetics. *Botanical Journal of the Linnean Society*. **169**(1): 5-40.

Ciocârlan, V. 2009. Flora ilustrată a României: Pteridophyta et Spermatophyta [Illustrated Flora of Romania: Pteridophytes and Spermatophytes]. Editura Ceres, Bucureşti.

Ciocârlan, V., and Sârbu, I. 2001. Taxonomia, variabilitatea şi răspândirea unor specii de *Astragalus* L. în Flora României. [Taxonomy, variability, and distribution of certain *Astragalus* L. species in the Flora of Romania]. *Buletinul Grădinii Botanice Iaşi*. **10**: 59-61.

Clement, M., Posada, D., and Crandall, K. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology*. **9**: 1657-1659.

Dong, W., Liu, J., Yu, J., Wang, L., and Zhou, S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS ONE*. **7**(4): e35071.

Downie, S. R., Katz-Downie, D. S., and Watson, M. F. 2000. A phylogeny of the flowering plant family Apiaceae based on chloroplast DNA *rpl*16 and *rpo*C1 intron sequences: towards a suprageneric classification of subfamily Apioideae. *American Journal of Botany*. **87**(2): 273-292.

Doyle, J. J., Doyle, J. L., Rauscher, J. T., and Brown, A. H. D. 2004. Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). *New Phytologist*. **161**(1): 121-132.

Drew, B. T., and Sytsma, K. J. 2011. Testing the monophyly and placement of *Lepechinia* in the tribe Mentheae (Lamiaceae). *Systematic Botany*. **36**(4): 1038-1049.

Drummond, A. J., and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*. **7**: 214-222.

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*.

Farris, J. S., Källersjö, M., Kluge, A. G., and Bult, C. 1994. Testing significance of incongruence. *Cladistics*. **10**(3): 315-319.

Ghahremani-Nejad, F. 2004. The sections of *Astragalus* L. with bifurcating hairs in Iran. *Turkish Journal of Botany*. **28**: 101-117.

Goncharov, N. F., Borisova, A. G., Gorshkova, S. G., Popov, M. G., and Vasilchenko, I. T. 1946. *Astragalus*. *In* Flora of the U.S.S.R. *Edited by* Komarov, V. L., and Shishkin, B. K. Izdatel'stvo Akademii Nauk S.S.S.R., Moscow, Leningrad. pp. 1-918. [Translated from Russian by Israel Prog. Sci. Trans., Jerusalem, 1965.].

Grant, V. 1981. Plant speciation. Columbia University Press, New York.

Gurushidze, M., Fritsch, R. M., and Blattner, F. R. 2010. Species-level phylogeny of *Allium* subgenus *Melanocrommyum*: Incomplete lineage sorting, hybridization and trnF gene duplication. *Taxon*. **59**(3): 829-840.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. **41**: 95-98.

Hewitt, G. H. 2000. The genetic legacy of quaternary ice ages. *Nature*. **405**: 907-913.

Hillis, D., Moritz, C., Porter, C., and Baker, R. 1991. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. *Science*. **251**(4991): 308-310.

Huson, D. H., and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*. **23**: 254–267.

Javanmardi, F., Kazempour Osaloo, S., Maassoumi, A. A., and Nejadsattrai, T. 2012. Molecular phylogeny of *Astragalus* section *Alopecuroidei* (Fabaceae) and its allies based on nrDNA ITS and three cpDNAs, matK, trnT-trnY and trnH-psbA sequences. *Biochemical Systematics and Ecology*. **45**: 171-178.

Jávorka, S. 1916. *Astragalus peterfii* Jáv. spec. nova. *Schedae ad Flora Hungaricae Exsiccatae*. **4**: 38-40.

Jorgensen, J. L., Stehlik, I., Brochmann, C., and Conti, E. 2003. Implications of ITS sequences and RAPD markers for the taxonomy and biogeography of the *Oxytropis campestris* and *O. arctica* (Fabaceae) complexes in Alaska. *American Journal of Botany*. **90**(10): 1470-1480.

Judd, W. S., Campbell, C. S., Kellog, E. A., Stevens, P. F., and Donoghue, M. J. 2008. Plant systematics: a phylogenetic approach. Sinauer Associates, Sunderland, MA.

Kazemi, M., Kazempour Osaloo, S., Maasoumi, A. A., and Pouyani, E. R. 2009. Molecular phylogeny of selected Old World *Astragalus* (Fabaceae): incongruence among chloroplst *trn*L-F, *ndh*F and nuclear ribosomal DNA ITS sequences. *Nordic Journal of Botany*. **27**(5): 425-436.

Kazempour Osaloo, S., Maasoumi, A. A., and Murakami, N. 2003. Molecular systematics of the genus *Astragalus* L. (Fabaceae): Phylogenetic analyses of nuclear ribosomal DNA internal transcribed spacers and chloroplast gene *ndh*F sequences. *Plant Systematics and Evolution*. **242**: 1-32.

-. 2005. Molecular systematics of the Old World *Astragalus* (Fabaceae) as inferred from nrDNA ITS sequence data. *Brittonia*. **57**(4): 367-381.

Kolarčik, V., Zozomová-Lihová, J., and Mártonfi, P. 2010. Systematics and evolutionary history of the Asterotricha group of the genus *Onosma* (Boraginaceae) in central and southern Europe inferred from AFLP and nrDNA ITS data. *Plant Systematics and Evolution*. **290**(1): 21-45.

Lavin, M., Herendeen, P. S., and Wojciechowski, M. F. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic Biology*. **54**(4): 575-594.

Ledingham, G. F. 1960. Chromosome numbers in *Astragalus* and *Oxytropis*. *Canadian Journal of Genetics and Cytology*. **2**(2): 119-128.

Ledingham, G. F., and Rever, B. M. 1963. Chromosome numbers of some Southwest Asian species of *Astragalus* and *Oxytropis* (Leguminosae). *Canadian Journal of Genetics and Cytology*. **5**(1): 18-32.

Liston, A. 1992. Isozyme systematics of *Astragalus* L. sect. *Leptocarpi* subsect. *Californici* (Fabaceae). *Systematic Botany*. **17**(3): 367-379.

Liu, Q., Ge, S., Tang, H., Zhang, X., Zhu, G., and Lu, B.-R. 2006. Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ribosomal internal transcribed spacer and chloroplast trnL-F sequences. *New Phytologist*. **170**(2): 411-420.

Lock, J. M., and Schrire, B. D. 2005. Galegeae. *In* Legumes of the World. *Edited by* Lewis, G., Schrire, B., Mackinder, B., and Lock, M. Kew, Richmond, Surrey. pp. 475-487.

Májovský, J., Uhríková, A., Javorčíková, D., Mičieta, K., Králik, E., Dúbravcová, Z., Feráková, V., Murín, A., Černušáková, D., Hindáková, M., Schwarzová, T., and Záborský, J. 2000. Prvý doplnok karyotaxonomického prehľadu flóry Slovenska [Karytaxonomical review of the flora of Slovakia]. *Acta Facultatis Rerum Naturalium Universitatis Comenianae, Botanica*. **Suppl. 1**: 1-127.

Masoud, S., Zarre, S., and Ismeilzadeh, J. 2009. New chromosome number reports in tragacanthic *Astragalus* species. *Caryologia*. **62**(1): 30-36.

Neubig, K., Whitten, W., Carlsward, B., Blanco, M., Endara, L., Williams, N., and Moore, M. 2008. Phylogenetic utility of *ycf*1 in orchids: a plastid gene more variable than *mat*K. *Plant Systematics and Evolution*. **277**(1): 75-84.

Neubig, K. M., Whitten, W. M., Williams, N. H., Blanco, M. A., Endara, L., Burleigh, J. G., Silvera, K., Cushman, J. C., and Chase, M. W. 2012. Generic recircumscriptions of *Oncidiinae* (Orchidaceae: Cymbidieae) based on maximum likelihood analysis of combined DNA datasets. *Botanical Journal of the Linnean Society*. **168**(2): 117-146.

Nieto Feliner, G., and Roselló, J. A. 2007. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution*. **44**(2): 911-919.

Nylander, J. A. A. (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. In.

O'Kane, S., Schaal, B., and Al-Shehbaz, I. 1996. The origins of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences. *Systematic Botany*. **21**(4): 559-566.

Page, R. D. M. 1996. Treeview: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*. **12**(4): 357–358.

Pânzaru, P. 2006. Note sistematiche e distributive relative a *Astragalus pastellianus* Pollini e a *A. peterfii* Jáv. (Leguminosae) [Taxonomic and distributive notes on *Astragalus pastellianus* Pollini and *A. peterfii* Jáv. (Leguminosae)]. *Bollettino del Museo Regionale di Scienze Naturali di Torino*. **23**: 721-727.

Parks, M., Cronn, R., and Liston, A. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology*. **7**(1): 84.

Pavlova, D., and Kozhuharov, S. 1993. Chromosome numbers of Bulgarian angiosperms. *Fitologija*. **44**: 75-76.

Philippov, E. G., Kulikov, P. V., and Knjasev, M. S. 2008. Chisla khromosom vidov *Astragalus* i *Hedysarum* (Fabaceae) flory Rossii [Chromosome numbers of *Astragalus* and *Hedysarum* species of the Russian flora]. *Botanicheskij Zhurnal*. **93**(10): 1614-1619.

Pirie, M. D., Vargas, M. P. B., Botermans, M., Bakker, F. T., and Chatrou, L. W. 2007. Ancient paralogy in the cpDNA *trn*L-F region in Annonaceae: implications for plant molecular systematics. *American Journal of Botany*. **94**(6): 1003-1016.

Podlech, D. 1990. Die Typifizierung der altweltlichen Sektionender Gattung *Astragalus* L. (Leguminosae). *Mitteilungen der Botanischen Staatssammlung Mü nchen* **29**: 461-494.

-. 2008. The genus *Astragalus* L. (Fabaceae) in Europe with exclusion of the former Soviet Union. *Feddes Repertorium*. **119**(5-6): 310-387.

-. 2011. Thesaurus Astragalorum: Index of all taxa described within the genus *Astragalus* L. and other genera but belonging to the genus *Astragalus*. Taxa of the Old World and related taxa of the New World, Ludwig-Maximilians-Universität München.

Posada, D. (2006) Collapse: Describing haplotypes from sequence alignments, Version 1.2. Available from: http://darwin.uvigo.es/software/collapse.html. In.

Pretel Martínez, A. 1974. IOPB. Chromosome number reports. XLVI. *Taxon*. **23**(5/6): 804-805.

Ranjbar, M. 2004. *Astragalus* sect. *Dissitiflori* (Fabaceae) in Iran. *Nordic Journal of Botany*. **24**(5): 523-531.

Ranjbar, M., and Karamian, R. 2002. Taxonomic study of *Astragalus* sect *Erioceras* (Fabaceae) in Iran. Additional notes and key to the species. *Nordic Journal of Botany*. **22**(6): 713-717.

Riahi, M., Zarre, S., Maassoumi, A., Kazempour Osaloo, S., and Wojciechowski, M. F. 2011. Towards a phylogeny for *Astragalus* section *Caprini* (Fabaceae) and its allies based on nuclear and plastid DNA sequences. *Plant Systematics and Evolution*. **293**(1): 119-133.

Richardson, J. E., Fay, M. F., Cronk, Q. C. B., Bowman, D., and Chase, M. W. 2000. A phylogenetic analysis of Rhamnaceae using *rbc*L and *trn*L-F plastid DNA sequences. *American Journal of Botany*. **87**(9): 1309-1324.

Rieseberg, L. H., and Brunsfeld, S. J. 1992. Molecular evidence and plant introgression. *In* Molecular Systematics of Plants. *Edited by* Soltis, P. S., Soltis, D. E., and Doyle, J. J. Chapman & Hall, New York. pp. 151–176.

Riggins, C. W., and Seigler, D. S. 2012. The genus *Artemisia* (Asteraceae: Anthemideae) at a continental crossroads: Molecular insights into migrations, disjunctions, and reticulations among Old and New World species from a Beringian perspective. *Molecular Phylogenetics and Evolution*. **64**(3): 471-490.

Sang, T., Crawford, D. J., and Stuessy, T. F. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. *Proceedings of the National Academy of Sciences*. **92**(15): 6813-6817.

Scherson, R. A., Choi, H.-K., Cook, D. R., and Sanderson, M. J. 2005. Phylogenetics of New World *Astragalus*: Screening of novel nuclear loci for the reconstruction of phylogenies at low taxonomic levels. *Brittonia*. **57**(4): 354-366.

Scherson, R. A., Vidal, R., and Sanderson, M. J. 2008. Phylogeny, biogeography, and rates of diversification of New World *Astragalus* (Leguminosae) with an emphasis on South American radiations. *American Journal of Botany*. **95**(8): 1030-1039.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., Siripun, K. C., Winder, C. T., Schilling, E. E., and Small, R. L. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*. **92**(1): 142-166.

Shaw, J., Lickey, E. B., Schilling, E. E., and Small, R. L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*. **94**(3): 275-288.

Silvestro, D., and Michalak, I. 2011. RaxmlGUI: a graphical front-end for RAxML. *Organisms, Diversity & Evolution*. http://dx.doi.org/10.1007/s13127-13011-10056-13120.

Soltis, P. S., and Soltis, D. E. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology*. **60**: 561 – 588.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. **22**(21): 2688-2690.

Stamatakis, A., Hoover, P., and Rougemont, J. 2008. A rapid Bootstrap algorithm for the RAxML web servers. *Systematic Biology*. **57**(5): 758-771.

Stuessy, T. F. 2008. Plant taxonomy: The systematic evaluation of comparative data. Columbia University Press, New York.

Swofford, D. L. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods). Sinauer Associates, Sunderland.

Sytin, A. K. 2009. Milk-vetches (*Astragalus* L., Fabaceae) of Eastern Europe and the Caucasus: systematics, geography, evolution. D.Sc. thesis, Institute of Botany, Russian Academy of Sciences, Sanct-Petersburg, Russia.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*. **28**(10): 2731-2739.

Thompson, J. D. 2005. Plant evolution in the Mediterranean. Oxford University Press, Oxford.

Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*. **34**: 275–305.

Walker, J. D., and Geissman, J. W. 2009. Geologic time scale. *Geological Society of America, GSA Today*. **19**: 60-61.

Wendel, J. F., and Doyle, J. J. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. *In* Molecular systematics of plants II: DNA sequencing. *Edited by* Soltis, D. E., Soltis, P. S., and Doyle, J. J. Kluwer, Dordrecht, the Netherlands. pp. 265–296.

Weng, M.-L., Ruhlman, T. A., Gibby, M., and Jansen, R. K. 2012. Phylogeny, rate variation, and genome size evolution of *Pelargonium* (Geraniaceae). *Molecular Phylogenetics and Evolution*. **64**(3): 654-670.

White, T. J., Bruns, T., Lee, S., and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* PCR Protocols: A Guide to Methods and Applications. *Edited by* Innis, M. A., Gelfand, D. H., Shinsky, J. J., and White, T. J. Academic Press, San Diego. pp. 315-322.

Willis, K. J. 1996. Where did all the flowers go? The fate of temperate European flora during glacial periods. *Endeavour*. **20**: 110-114.

Willyard, A., Cronn, R., and Liston, A. 2009. Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Molecular Phylogenetics and Evolution*. **52**(2): 498-511.

Wojciechowski, M. F. 2005. *Astragalus* (Fabaceae): A molecular phylogenetic perspective. *Brittonia*. **57**(4): 382-396.

Wojciechowski, M. F., Sanderson, M. J., Baldwin, B. G., and Donoghue, M. J. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): evidence from nuclear ribosomal DNA internal transcribed spacer sequences. *American Journal of Botany*. **80**(6): 711-722.

Wojciechowski, M. F., Sanderson, M. J., and Hu, J. 1999. Evidence on monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trn*L intron data. *Systematic Botany*. **24**(3): 409-437.

Xu, B., Wu, N., Gao, X.-F., and Zhang, L.-B. 2012. Analysis of DNA sequences of six chloroplast and nuclear genes suggests incongruence, introgression, and incomplete lineage sorting in the evolution of *Lespedeza* (Fabaceae). *Molecular Phylogenetics and Evolution*. **62**(1): 346-358.

Zhang, M., Kang, Y., Zhou, L., and Podlech, D. 2009. Phylogenetic origin of *Phyllolobium* with a further implication for diversification of *Astragalus* in China. *Journal of Integrative Plant Biology*. **51**(9): 889-899.

Zimmer, E. A., Martin, S. L., Beverley, S. M., Kan, Y. W., and Wilson, A. C. 1980. Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. *Proceedings of the National Academy of Sciences*. **77**(4): 2158-2162.

Zimmer, E. A., and Wen, J. 2012. Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular Phylogenetics and Evolution*. **65**(2): 774-785.

Zippel, E., and Wilhalm, T. 2009. Origin and relationships of *Astragalus vesicarius* subsp. *pastellianus* (Fabaceae) from the Vinschgau Valley (Val Venosta, Italy). *Gredleriana*. **9**: 119-134.

Zuccon, A., and Zuccon, D. 2006. MrEnt v2.2. Program distributed by the authors. Department of Vertebrate Zoology & Molecular Systematics Laboratory, Swedish Museum of Natural History, Stockholm.

Keywords: *Astragalus*, *Dissitiflori*, Erioceras, phylogeny, reticulation, *ycf*1