BioOne COMPLETE

The petD group II intron as a species level marker: utility for tree inference and species identification in the diverse genus Campanula (Campanulaceae)

Authors: Borsch, Thomas, Korotkova, Nadja, Raus, Thomas, Lobin, Wolfram, and Löhne, Cornelia

THOMAS BORSCH[1*,2], NADJA KOROTKOVA[3], THOMAS RAUS[1], WOLFRAM LOBIN[4] & CORNELIA LÖHNE[1]

The *petD* group II intron as a species level marker: utility for tree inference and species identification in the diverse genus *Campanula (Campanulaceae)*

Abstract

Borsch T., Korotkova N., Raus T., Lobin W. & Löhne C.: The *petD* group II intron as a species level marker: utility for tree inference and species identification in the diverse genus *Campanula (Campanulaceae)*. – Willdenowia 39: 7-33. – Online ISSN 1868-6397; © 2009 BGBM Berlin-Dahlem. doi:10.3372/wi.39.39101 (available via http://dx.doi.org/)

Chloroplast introns have a high potential as tools for phylogeny inference and DNA barcoding. This study examines the molecular evolution of the *petD* group II intron in *Campanulaceae* based on a sequence data set of 114 ingroup taxa. Three small mutational hotspots had to be excluded from phylogenetic analysis, the two most variable being located in the D4 loop (domain IV). A $(GT)_{4-7}$ microsatellite in domain II is conserved at species level but of limited phylogenetic use due to unclear homology of individual repeat units. Sequences of the *petD* group II intron depict *Cyphioideae, Lobelioideae* and *Campanuloideae* as major *Campanulaceae* clades. Core *Campanuloideae* comprise two major radiations of *Campanula* species: a *Musschia* clade (including *C. lactiflora*) and a *Jasione* clade. *Campanula* is highly paraphyletic to a number of smaller genera such as *Azorina, Michauxia* and *Edraianthus*. The closed-tubular flowered taxa (*Phyteuma* and allies) are resolved sister to *C. persicifolia*. Within core campanuloids *petD* sequences identify 90 % of the taxon samples included in this study. Considering the ease of amplification and sequencing, and its high information content, the *petD* intron appears to be a good candidate in a two-tailed approach integrating molecular phylogenetics and species identification in the needed sampling of all core *Campanuloideae* species.

Additional key words: chloroplast genome, phylogenetic structure, molecular evolution, Eurasia, endemics, DNA barcoding

1. Introduction

One of the greatest challenges in plant biodiversity research is to generate robust phylogenetic frameworks for the many species-rich genera as a prerequisite for reliable inference of character evolution, historical biogeography and to examine modes of speciation in time and space. Reconstructing phylogenetic relationships in large and species-rich genera as well as analysing character evolution and biogeographic patterns in these genera require the inclusion of many taxa. Typically, the need for extensive taxon sampling limits the amount of sequence data that can be generated per individual taxon. As a consequence, efficient markers are needed that provide a maximum of phylogenetic signal per base sequenced.

In angiosperms, chloroplast spacers and introns and also the rapidly evolving *matK* gene were used mostly for species level phylogenetic inference (Borsch & Quandt 2009). The group I intron in *trnL* and the *trnL-F* spacer (together the "*trnL-F* region") were proposed early on as markers that can be amplified and sequenced using universal primers (Taberlet & al. 1991) and are now the most frequently used non-coding plastid regions (Quandt & Stech 2003; Shaw & al. 2007). Other frequently used spacers are those between *atpB* and

---

1 Botanic Garten and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Königin-Luise-Straße 6-8, D-14195 Berlin, Germany; *e-mail: t.borsch@bgbm.org (author for correspondence).
2 Institute for Biology, Freie Universität Berlin, Altensteinstraße 6, D-14195 Berlin, Germany.
3 Nees Institute for Biodiversity of Plants, Friedrich-Wilhelms-Universität Bonn, Meckenheimer Allee 170, D-53115 Bonn, Germany.
4 Bonn University Botanical Gardens, Meckenheimer Allee 171, D-53115 Bonn, Germany.

Fig. 1. Floral diversity in the *Campanuloideae* – A: *Campanula barbata* L. *(Borsch 3842),* a member of the *C. latifolia* clade; B: *Phyteuma scheuchzeri* All. *(Borsch 3853),* a member of the *Phyteuma* clade with closed-tubular flowers; C: *C. rotundifolia* agg. *(Borsch 3847),* a member of the *C. rotundifolia* clade with typical bell-shaped flowers; D: *C. versicolor* Andrews *(Korotkova 91),* a member of the *C. rotundifolia* clade with a deeply split corolla; E: *Petromarula pinnata* (L.) A. DC. *(Korotkova 95),* a member of the *Phyteuma* clade with recurved, free corolla lobes; F: *Cyclocodon lancifolium* Kurz *(Neumann s.n.),* a member of the *Canarina* clade. – Species illustrated in A to E belong to the core *Campanuloideae*.

rbcL (e.g., Ehrendorfer & al. 1994; Manen & al. 1994), between *psbA* and *trnH* (Sang & al. 1997; Hamilton 1999) and between *trnS* and *trnG* (Hamilton 1999). More recently, group II introns were discovered as further powerful tools (e.g., Kelchner 2002), with the two

introns in *rpl*16 and *rps*16 currently being the most widely used. Considering that the greatest variability of plastid genomic regions between closely related taxa is found in the non-coding parts of the chloroplast genome's large and small single copy regions (LSC and

SSC), more and more spacers and introns from that part are used in plant evolutionary studies every year. Small & al. (1998) started a series of studies comparing the variability of sequences from 3-taxon sets of closely related species across angiosperms and compared so-called PIC values (parsimony informative characters) for many spacers and introns. More recently, such comparisons of variability are also based on completely sequenced genomes (e.g., Timme & al. 2007).

On the other hand, it is obvious that there are considerable differences in phylogenetic utility (corresponding to differences in phylogenetic structure) of genomic regions. These are determined by specific mutational dynamics of a genomic region governing the accumulation of character state changes during the region's evolutionary history. A comparison of sequence data sets from *rbcL, matK* and the non-coding parts of the *trnT-trnF* region (Müller & al. 2006) showed that phylogenetic structure is not only determined by the number of potentially informative sites (i.e., the amount of information) but also by the distribution of mutations across the tree (i.e., quality of information). Sequences of *trnT-trnF* performed significantly better than *matK* in a basal angiosperm taxon set. Consequently, phylogenetic structure in sequence data sets cannot be easily deduced from PICs in a small sequence set (Timme & al. 2007), but is determined by molecular evolutionary patterns, which again are strongly influenced by structural and functional constraints in a genomic region. Searching highly variable markers with the best phylogenetic structure is therefore a worthwhile approach (Borsch & Quandt 2009). In addition, there can of course be practical limitations, since not all genomic regions are easily amplified for a wide range of taxa (such as the *trnT-trnL* spacer; Borsch & Quandt 2009) or frequently cause sequencing problems due to long strands of A's or T's (>10 nt, such as the *psbA-trnH* spacer; Devey & al. 2009).

However, resolution and internal statistical support of the recovered trees are not sufficient in many cases. The trend therefore is to combine more regions for increased quality of phylogenetic hypotheses. Most current studies include at least two regions in combined data sets, and multiple data sets combining four to five chloroplast regions are becoming standard (e.g., Barfuss & al. 2005; Kocyan & al. 2007; Löhne & al. 2007; Kårehed & al. 2008) to infer a robust plastid tree. There are many cases where the combination of many noncoding and rapidly evolving chloroplast regions led to considerably improved tree resolution and support within and among genera of various angiosperm lineages (e.g., Barfuss & al. 2002; Tesfaye & al. 2007; Löhne & al. 2007; Kårehed & al. 2008). Tree reconstruction in *Campanulaceae* and especially in *Campanula* L. has so far only yielded partly supported trees, including large polytomies (nrITS: Eddie & al. 2003; Roquet & al. 2008; *trnL-F:* Roquet & al. 2008). Most recently combination of *rbcL, atpB* and *matK* (approx.

4200 nt) yielded improved but still not fully resolved trees (Cellinese & al. 2009).

The *petD* gene is part of the *petB* operon consisting of five genes, *psbB, psbT, psbH, petB* and *petD,* in the LSC region of the chloroplast genome (Westhoff & Hermann 1988). The group II intron in *petD* is located in the upstream part of the gene, following an 8 nt 5' exon. Like other group II introns it is composed of six domains with conserved helical elements that are arranged around a central wheel, whereas the distal parts of especially domains I and IV are more variable. Universal primers for amplifying the *petD* intron have been designed by Löhne & Borsch (2005). Since then these primers have been successfully applied to inferring deep level relationships in eudicots (Worberg & al. 2007), rosids (Worberg & al. 2009) and asterids (Salomo & al., unpubl. data) as well as in *Nymphaeales* (Löhne & al. 2007) and *Malpighiales* (Korotkova & al. 2009). Compared to other chloroplast regions (e.g., *trnK/matK, trnT-F, rpl16, atpB-rbcL*) the amplification of *petD* was always easy and yielded large amounts of PCR products. The alignment always was straightforward, with only small and well defined mutational hotspots. As a phylogenetic marker *petD* performed well in all the studies carried out so far. More recently, Kårehed & al. (2008) and Groeninckx & al. (2009) showed *petD* intron sequences in *Rubiaceae* to contain much better phylogenetic signal than all other previously used markers.

It was thus promising to ask for variability and performance of *petD* intron sequences in phylogenetic analyses of a species-rich lineage. The *Campanulaceae* (the bell flower family) and the genus *Campanula* (Fig. 1) are diverse in northern temperate and Mediterranean regions. About 150 of the 300-400 species of the genus *Campanula* occur in Europe (Fedorov & Kovanda 1976). More than 260 plus many infraspecific taxa occur in the Mediterranean area (Greuter & al. 1984), of which more than 80 % are endemic to that area. Hotspots of endemism are found in the E Mediterranean (Phitos 1964, 1965; Carlström 1986), the Balkan region (Kovačić 2004), the Caucasus (e.g., Gagnidze 2005) and Turkey (Damboldt 1978). Moreover, generic concepts in *Campanulaceae* vary considerably, with between 40 and 85 recognized genera (Kolakovsky 1987, 1994; Eddie & al. 2003; Lammers 2007), many of the smaller ones being European or W Asian endemics. Phylogenies published so far agree on high paraphyly or polyphyly of the genus *Campanula,* resulting in the need for additional dense taxon sampling. At the same time there are many taxonomically difficult groups of morphologically closely allied taxa with unclear status (see Eddie & al. 2003). In various lineages of *Campanula* there is high phenotypic plasticity and there are polyploid series (Podlech & Damboldt 1964; Kovanda 1970a, b, 1977). As a consequence, taxonomic treatments at the species level often differ considerably among different authors and for different countries, causing serious problems for

assessing distribution and conservation status of the species. A striking example is the *C. rotundifolia* group (Podlech 1965; Fedorov & Kovanda 1976). DNA sequence data may complement morphology and help to back up taxonomic decisions and conservation measures. With a perspective on the conservation of the palearctic flora, *Campanula* s.l. is therefore an important genus in need of considerable research in order to effectively plan conservation measures.

The *Campanulaceae* belong to the eudicot order *Asterales* and contain approximately 2300 species (Lammers 2007). Classification systems either treat the closely related *Lobeliaceae* as separate family or include a subfamily *Lobelioideae* into a broadly defined family *Campanulaceae* (e.g., Cronquist 1988; APG II 2003; Lammers 2007; followed here). In other classification systems, *Lobeliaceae* are considered as an own family close to *Campanulaceae* (Takhtajan 1997). Nevertheless, phylogenetic relationships within the broad *Campanulaceae* alliance were never tested using a dense taxon sampling from all putative major lineages of the bell flower family, but rather focused on *Campanuloideae* (Eddie & al. 2003; Roquet & al. 2008) or *Lobelioideae* (Antonelli 2008), using the respective other lineage as outgroup. In analyses of *Asterales* based on nuclear ITS sequences, Eddie & al. (2003) first provided evidence that *Campanula* is not monophyletic. Due to its high variability, the nrITS region has been widely used for phylogeny inference at the species and genus levels. However, the backbone of ITS trees in *Campanulaceae* is poorly resolved and many nodes did not gain good statistical support. Pollen (Dunbar 1975 a,b) and seed (Belyaev 1984 a,b) characters are informative in *Campanulaceae* and were used among other morphological data in a phylogenetic analysis by Shulkina & al. (2003). Larger taxon sets were further generated of the *trnL-F* region (Roquet & al. 2008) and the *rbcL*+*atpB* +*matK* genes (Cellinese & al. 2009; hereafter referred to as 3-gene analysis or data set). However, taxon sampling differs considerably between the two last-named studies, and major parts of the topology are not receiving high statistical support, even in the 3-gene analysis. Thus, there is a need for sampling more taxa and more characters, ideally based on molecular markers chosen for high phylogenetic signal.

Aims of this study were (1) to generate a well sampled *petD* group II intron data set for the genus *Campanula* and the *Campanulaceae* and to infer phylogenetic hypotheses based on both substitutions and microstructural mutations; (2) to calculate a secondary structure of the *petD* intron for *Campanula* and to evaluate mutational hotspots in an accurate structural context; (3) to evaluate the phylogenetic utility of *petD* sequences for reconstructing species level phylogenies and the utility of *petD* sequences for species identification in *Campanula* s.l.

2. Material and methods

Taxon sampling and plant material — Sequences of the *petD* intron and the *petB-petD* intergenic spacer were obtained for 114 taxa currently classified as *Campanulaceae* incl. *Lobeliaceae* and one representative each of *Stylidiaceae* and *Pentaphragmataceae,* which are close relatives of *Campanulaceae* (Albach & al. 2001; Bremer & al. 2002; Lundberg & Bremer 2003). Plants for this project were collected, predominantly in Germany, Georgia, Italy and Slovakia. A further important source of material were the living collections and the herbarium of the Botanic Garden and Botanical Museum Berlin-Dahlem (B) and the living collection of the Bonn University Botanical Gardens. Four taxa were sampled as outgroups, one from *Asterales (Lactuca sativa* L.; *Asteraceae)* and three from more distantly related asterids *(Atropa belladonna* L. and *Nicotiana tabacum* L., *Solanaceae; Sphenoclea zeylanica* Gaertn., *Sphenocleaceae).* The sequence for *Sphenoclea* was generated in this study, whereas sequences for the other three taxa were available through completely sequenced chloroplast genomes. Vouchers have been deposited in the herbaria at Berlin (B) or Bonn (BONN). Detailed information on the sampled material is provided in Appendix 1.

DNA isolation, amplification and sequencing — Genomic DNA was isolated from silica dried leaf tissue or herbarium specimens using the modified CTAB method with triple extractions described by Borsch & al. (2003) or using the NucleoSpin Plant II Kit (Macherey-Nagel). The *petB-petD* spacer and the *petD* intron were co-amplified using the universal primers PIpetB1411F and PIpetD738R designed by Löhne & Borsch (2005). Amplification reactions contained 2 µl DNA template with a concentration of 10 ng/µl, 5 µl Taq buffer S (PeqLab [including 15 mM MgCl$_2$], 2 µl of each primer (20 pm/µl), 10 µl dNTPs (each 1.25 mM), and 1.5 units of Taq DNA Polymerase (PeqLab). Ultrapure H$_2$O was added to obtain a total volume of 50 µl. PCR amplifications were performed in a T3 Thermocycler (Biometra). Amplification conditions were as follows: an initial denaturation step of 1.5 min at 96 °C, followed by 0.5 min at 95 °C, 1 min at 50 °C, 1.5 min at 72 °C, for 34 cycles, and a final extension step of 20 min at 72 °C. The PCR products were stained with 100x Gelstar nucleic acid stain and electrophoresed on a 2 % agarose gel at 80 V for 2 hours. Fragments were excised from the gel and DNA was extracted and purified using the Qiaquick Gel Extraction Kit (QIAGEN) or the Gel/PCR DNA Fragment Extraction Kit (Avegene). The purified products were directly sequenced using the DCTS Quick Start Kit (Beckman Coulter) and run on an automated capillary sequencer (CEQ 8000 Genetic Analysis System, Beckman Coulter) or sequenced via Macrogen Inc. (Seoul, South Korea). Pherograms were

edited manually with PhyDe992 and later versions (Müller & al. 2005+).

Sequence alignment, coding of length mutational events — Sequences were aligned manually using PhyDE0992 and later versions (Müller & al. 2005+). Rules for the alignment of non-coding regions as outlined by Borsch & al. (2003) and Löhne & Borsch (2005) are followed here. Regions of uncertain homology (hotspots sensu Borsch & al. 2003) were excluded from the analysis. Indels were coded according to the Simple Indel Coding method (Simmons & Ochoterena 2000) using SeqState 1.40 (Müller 2005a).

Phylogenetic analyses — A parsimony ratchet was performed using PRAP (Müller 2004a). Ratchet settings were 200 iterations with 25 % of the positions randomly upweighted (weight = 2) during each replicate and 10 random addition cycles. The number of steps for each tree and the consistency, retention and rescaled consistency indices (CI, RI and RC) were calculated using PAUP* v. 4.0b10 (Swofford 1998). Support for the nodes found by the parsimony ratchet was inferred using jackknifing with the optimal jackknife parameters described by Müller (2005b). A total number of 10 000 jackknife replicates was performed using the TBR branch swapping algorithm with 36.788 % of characters deleted in each replicate. One tree was held during each replicate.

Bayesian Inference (BI) was carried out using MrBayes 3.1 (Huelsenbeck & Ronquist 2001). Nucleotide substitution models for the data set were evaluated using Modeltest 3.7 (Posada & Crandall 1998) with spacer and intron sequences analysed separately. The hierarchical likelihood ratio test (hLRT) suggested the GTR+Γ+I model. Four simultaneous runs of Metropolis-coupled Markov Chain Monte Carlo (MCMCMC) analyses, each with four parallel chains, were performed for 1 million generations, saving one tree every 100th generation, starting with a random tree. Other MCMC parameters were left with the program's default settings. Burn-in was set at generation 50 000. The remaining trees were summarised in a majority rule consensus tree.

Inference of RNA secondary structure — The sequence of *Campanula trachelium* L. was chosen to represent a typical *petD* intron sequence of *Campanula* without large indels or other deviant features. Due to the large size of the intron, current available algorithms for RNA folding are not able to fold the entire intron sequence in one step, since too many alternative foldings are possible (Mathews & al. 2006). To ensure that the adequate secondary structure was inferred, structural partitions of the intron, i.e., domains, subdomains, were first identified based on the annotated *petD* intron alignment of Michel & al. (1989). Then, each of the domains was folded separately following Kelchner (2002). Addi-tionally, constraints for the two exon binding sites and the branch-point "A" were defined to leave these sequence parts single-stranded. Foldings were done using RNAstructure 4.5 (Mathews & al. 1996+) that is based on the algorithm of Mathews & al. (2004). The intron structure was then drawn using RnaViz 2.0 (Rijk & al. 2003).

Calculation of substitutional rates — Relative rate tests were performed using GRate version 1.0 (Müller 2004b). This software uses the DNA substitution models implemented in phylogenetic analyses and allows to compare rates between pre-defined groups of taxa. All major clades of *Campanulaceae* s.l. found with parsimony and Bayesian tree inference in this study were defined as groups for the comparison of molecular rates. *Atropa bella-donna* was set as outgroup and *Nicotiana tabacum* as reference taxon.

3. Results

The *petD* intron and the upstream *petB-petD* spacer were easily amplified and sequenced for DNAs isolated from both silica gel dried and herbarium material. Sequencing was not hindered by long stretches of As/Ts (chloroplast microsatellites) and single reads were covering the whole region when using an ABI3100 system. Therefore, the use of both amplification primers for sequencing generated a double coverage of the marker.

The *petD* intron ranged from 654-772 nt in *Campanulaceae* and the *petB-petD* spacer from 172-231 nt. The secondary structure of the *petD* intron calculated for *Campanula trachelium* is given in Fig. 2. Examples of the secondary structure of its domain II, representing different sizes of the $[GT]_n$ microsatellite, are shown in Fig. 3.

In the overall alignment three small hotspots (HS1-HS3) had to be excluded (Fig. 2, Appendix 3) from the phylogenetic analysis. The alignment can be downloaded (at www.eudicots.de) and the EMBL/GenBank accession numbers of the sequences are given in Appendix 1. The matrix comprised 1362 characters in total, of which 182 were variable but not informative and 372 were variable and parsimony informative. Maximum Parsimony reconstruction and Bayesian Inference yielded strict consensus and majority rule trees for *Campanulaceae* and *Campanula* (Fig. 4, 5) that were largely resolved and well supported except in the *C. latifolia* clade. Parsimony analysis based on substitutions only resulted in 350 shortest trees (1303 steps) with a CI = 0.629, a RI = 0.869 and a RC = 0.546. A list of postulated microstructural mutations is provided in Appendix 2. Simple sequence repeats accounted for most of these mutations, whereas inversions were not found. Addition of a matrix of 192 indels raised the number of variable but not informative characters to 299 and of variable and parsimony-informative characters to 437. From the combined matrix 229 trees of a length of 1525
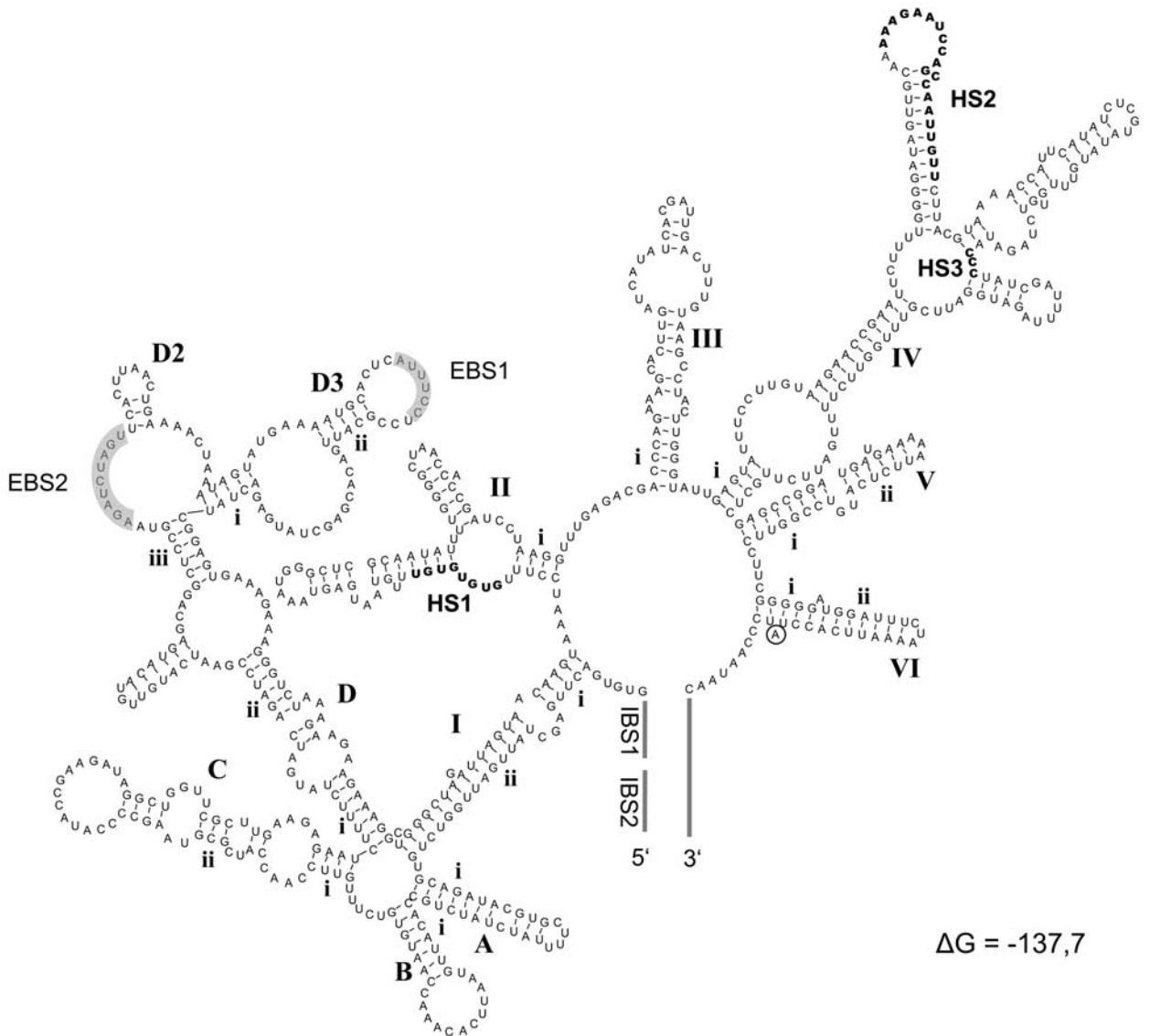
Fig. 2. Secondary structure of the *petD* group II intron from *Campanula trachelium*. – Nomenclature of structural elements follows Michel & al. (1989). Roman numerals I to VI highlight domains. Subdomains of domain I are labelled with letters A-D(2). Lower case roman numerals i to iii mark conserved helical elements. Domain IV contains the highlighted branch-point "A". EBS = "Exon Binding Site" and IBS = "Intron Binding Site".

steps and a CI = 0.656, a RI = 0.868 and a RC = 0.570 were recovered. The strict consensus is shown in Fig. 4. Bayesian Inference sampled 9500 trees from four chains of four independent runs employing a GTR+Γ+I model. The majority rule tree is shown in Fig. 5 as a phylogram.

The average relative substitutional rates in the *petD* intron data set calculated for all major *Campanulaceae* groups in the present study are given in Fig. 6.

4. Discussion

RNA secondary structure of the *petD* intron in *Campanula*

As an example for the core *Campanuloideae,* the secondary structure of the *petD* intron in *Campanula trachelium* was calculated (Fig. 2). The sequence of *C.*

trachelium showed comparatively few microstructural mutations and was therefore considered to be a representative example for the study group. The overall structure reflects the typical consensus structure of group II B introns with six domains arranged around a central wheel (Michel & al. 1989; Kelchner 2002). Domain structure of the *petD* intron in *Campanula* compares well in size and variability to the secondary structure of the intron in asterids in general (Salomo & al., unpubl. data) and other lineages of eudicots (e.g., *Malpighiales,* Korotkova & al. 2009). Domain I is the largest domain with a conserved helical structure and four subdomains. Subdomain D2 of domain I is extremely small in *Campanula* but can become very large, also due to the insertion of repetitive elements, in other eudicots (Korotkova & al. 2009). Relative to the *petD* intron in other asterids
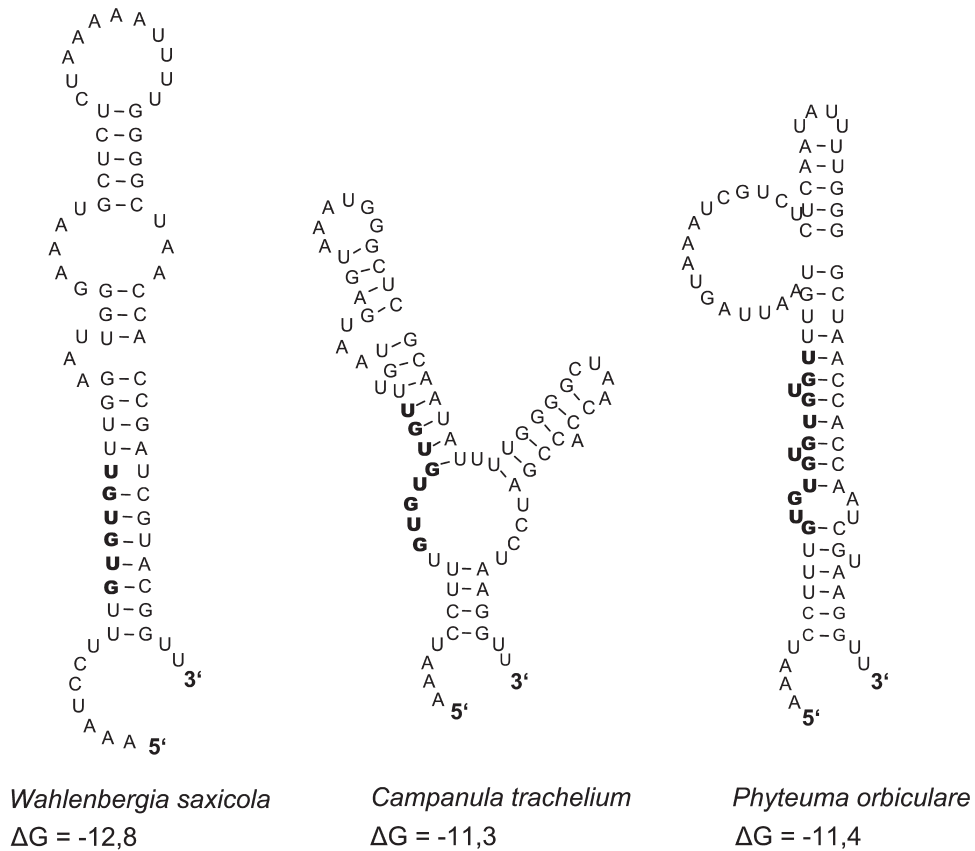
Fig. 3. Secondary structure of domain II of the *petD* group II intron in three species representing different sizes of the [GT]$_n$ microsatellite (hotspot 1). – *Wahlenbergia saxicola* A. DC. with three repeat units, *Campanula trachelium* with four repeat units and *Phyteuma orbiculare* L. with six repeat units.

and eudicots the complete domain II and the terminal stem-loop of domain IV deviate most, whereas domains V and VI are largely conserved.
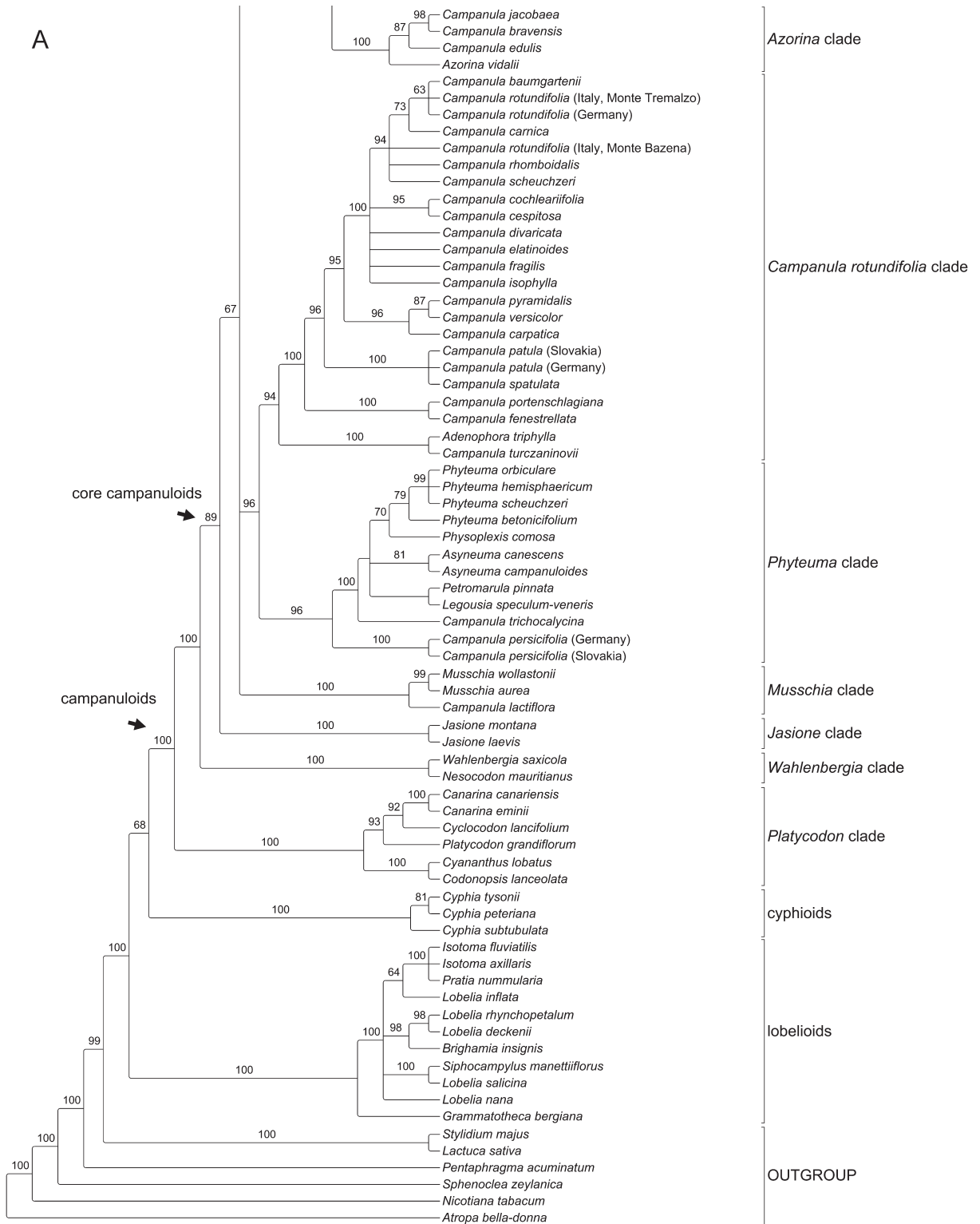
Patterns of variability in the *petD* group II intron

The *petD* intron and the upstream *petB*-*petD* spacer yielded a reliable alignment with the majority of micro-structural mutations being simple sequence repeats (SSRs; Appendix 2) that allowed unambiguous motif recognition. The sequences analysed in this study cover a much smaller spectrum of lineage diversity than all *petD* data sets previously published in angiosperms (Löhne & Borsch 2005; Worberg & al. 2007; Korotkova & al. 2009). Even though the distances between the sequences are much smaller, three mutational hotspots of unclear sequence homology across all *Campanulaceae* and the outgroups had to be excluded from the analysis (Fig. 2, Appendix 3). Compared to Löhne & Borsch (2005), who used a consensus model for group II introns (Michel & al. 1989) to annotate domain borders in the alignment and to map hotspots onto the secondary structure, this study allows a more precise examination of domain structure and of the exact location of mutational hotspots within the domains.

Compared to other rapidly evolving spacer and intron data sets the percentage of nucleotides in hot-

spots is extremely low (Appendix 3). On average hotspots comprise less than 4 % of nucleotides in the *petD* intron of *Campanulaceae*. One of the reasons is a distally scarcely extended subdomain D2 of domain I. In all *petD* intron secondary structures known so far this subdomain is usually AT-rich and length variable, as in most other asterids (Salomo & al., unpubl. data) and in *Malpighiales* (Korotkova & al. 2009).

Hotspot HS1 is a dinucleotide GT microsatellite (Fig. 3). It is also present in other asterids (Salomo & al., unpubl. data) and is constituted by three to six repeat units in *Campanulaceae*. Although homoplastic at the level of the family, repeat number is conserved within major clades. No variation within species or among closely allied taxa (e.g., *Campanula rotundifolia* L. and its allies) is observed. The *C. latifolia* and *Azorina* clades have a copy number of four, the *C. rotundifolia* clade of five, the *Phyteuma* clade also of five but raising to six in *Phyteuma* s.str. All other *Campanuloideae* also have five repeat units with the exception of *Wahlenbergia* (three). The conserved nature of this microsatellite may be caused by its position in a stem element of domain II where it is stabilised through complementary base-pairing. An incrementation of the repeat-unit copy number extends the length of the proximal helix in domain II and leads to a considerably altered distal part of

A

Azorina clade
- 98, 87 — *Campanula jacobaea*, *Campanula bravensis*
- 100 — *Campanula edulis*
- *Azorina vidalii*

Campanula rotundifolia clade

Phyteuma clade

Musschia clade

Jasione clade

Wahlenbergia clade

Platycodon clade

cyphioids

lobelioids

OUTGROUP

core campanuloids

campanuloids

this domain in *C. trachelium* (one of the taxa with four repeat units; Fig. 3). The structure shown is thermodynamically suboptimal with a ΔG value of -11.3, whereas the optimal folding (ΔG = -13.0) forms a second helix departing from the intron's central wheel (not in line with a group II intron consensus structure). High mutational activity in domain II appears to be present in core eudicots, as evidenced by this study and observations by Korotkova & al. (2009) and Salomo & al. (unpubl. data), but not in early branching angiosperms (Löhne & Borsch 2005). Domain II is the least important for correct intron splicing (Lehmann & Schmidt 2003), so that the mutational hotspot again correlates with an area of minimal evolutionary constraints.
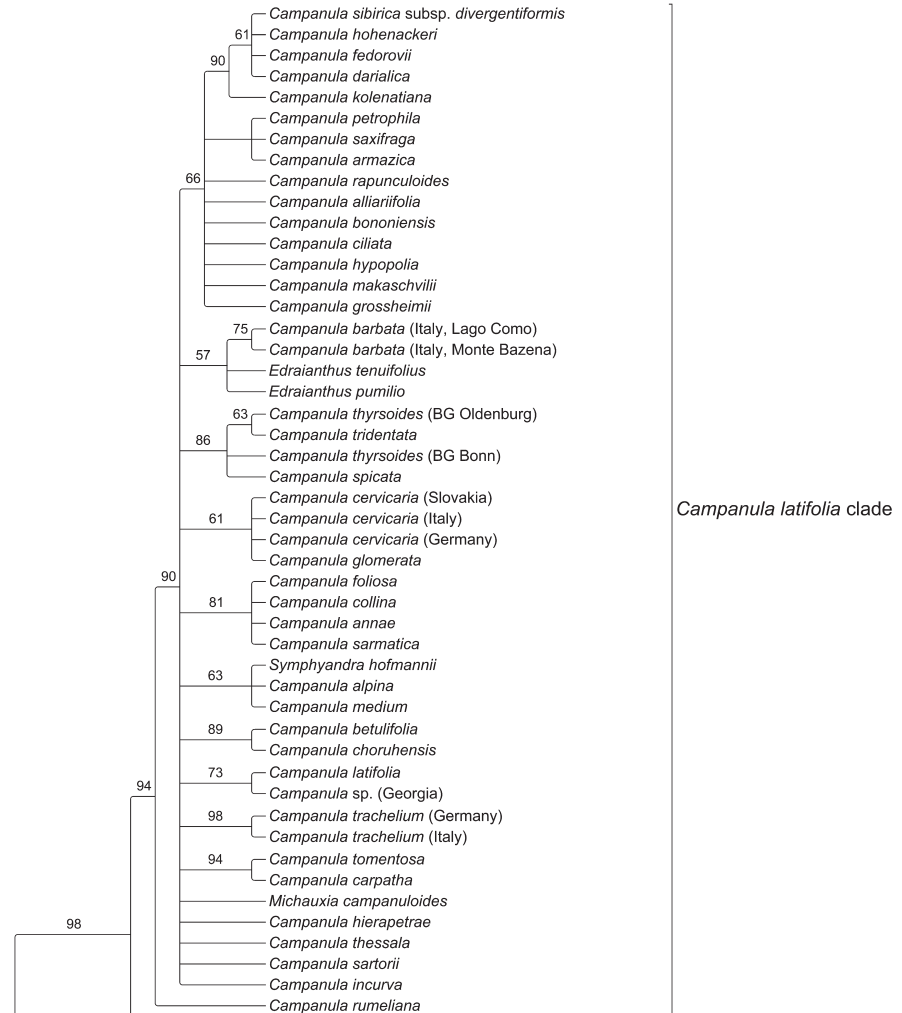
Fig. 4A-B. Maximum Parsimony tree of *Campanulaceae* (strict consensus of 229 shortest trees of 1525 steps) based on the *petD* data set including indels. – Jackknife values supporting individual nodes are shown above branches.

The other two hotspots are located in the D4 loop of domain IV. In *Campanulaceae* this loop is the largest and most variable terminal stem-loop. High D4 loop variability appears to be a general pattern of plant chloroplast introns (Kelchner 2002; Löhne & Borsch 2005; Watts & al. 2008; Korotkova & al. 2009), as evolutionary constraints are low because of the degradation of the maturase open reading frames in all introns except *trnK* (Toor & al. 2001).

Hotspot HS 2 is mostly generated by insertion/ deletion events that lead to different terminal loops (Fig. 2). However, within core *Campanuloideae* microstructural mutations are rather rare (a 4 nt deletion in *Symphyandra hofmannii* Pant. and a "AAAGAA" SSR in *C. rumeliana* (Hampe) Vatke) so that the respective sequence parts could be utilized in phylogenetic analyses limited to that clade. The same regards to HS 3 that consists of three nucleotides in *Campanuloideae* (Fig. 2), whereas it is length variable in *Cyphioideae, Lobelioideae* and the other *Asterales*.

Caused by their core structure, group II introns exhibit mosaic patterns of sequence conservation (Kelchner 2002; Löhne & Borsch 2005). Helical elements tend to be conserved, whereas stem-loop elements are usually less constrained and may be expected to contain more variable sites as well as insertions, deletions and inversions. In line with the *petD* intron consensus structure in asterids (Salomo & al., unpubl. data) there are several conserved helical elements labelled i to iii throughout all intron domains in *Campanulaceae* (Fig. 2). In particular, the structure of the large domain I is upheld by these helical elements. For microstructural mutations and, as a consequence, for our alignment this results in an evolutionary constrained core structure that still exhibits many variable sites.

The *petD* intron is a rather small group II intron in the chloroplast genome, especially compared to the on average about 25 % bigger *rpl*16 intron (Borsch & Quandt 2009). Length differences among different intron sequences were primarily attributed to smaller
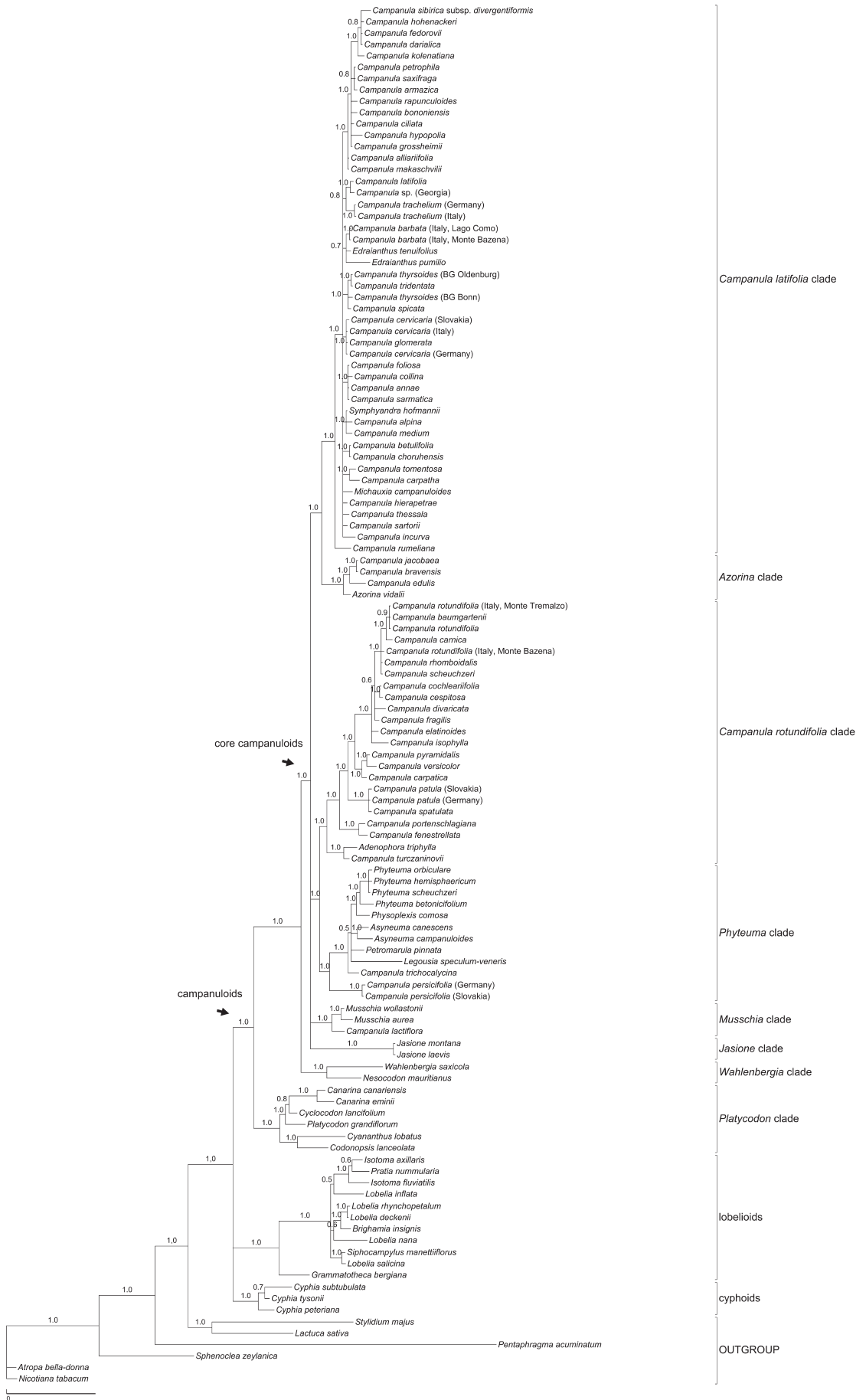
stem-loop elements in domains I (especially subdomain D2) and IV. In a comparison of intron variability in identical taxon sets in *Nymphaeales,* Löhne & al. (2007) found the *petD* intron to be the least variable, also yielding lower phylogenetic structure *R* in comparison to the introns in *rpl*16 and *trnK* and also to *trnT-trnF* (comprising two spacers and the group I intron in *trnL*). It is therefore rather surprising that this study finds high phylogenetic structure of the *petD* intron at species level in *Campanula* (see also below). If the group II intron secondary structure is a mosaic of more conserved helical elements and more variable loops and bulges, one would expect that point mutations are more frequent in the less constrained loops and bulges when less distant sequences are compared (e.g., those spanning different species and closely related genera). The *petD* intron in *Campanulaceae* will provide a good model to determine structural elements (stems, loops, bulges) based on precisely calculated secondary structures, and to evaluate the distribution of mutations in relation to the structure.

Circumscription of *Campanulaceae* and phylogenetic relationships within the *Campanulaceae* s.l.

Within *Asterales* the familes *Pentaphragmataceae, Stylidiaceae* and *Rousseaceae* have been inferred to be the closest relatives of the *Campanulaceae* clade (including *Lobeliaceae*). *Campanulaceae* and *Stylidiaceae* were sisters in the combined analysis of *rbcL*+*atpB*+18S+*ndhF* (Albach & al. 2001) but with only weak support (61 % JK). In a subsequent analysis using six chloroplast regions, Bremer & al. (2002) raised the confidence into this node to 82 % JK, whereas *Pentaphragmataceae* were hypothesised as sister to a *Campanulaceae-Stylidiaceae* clade. Lundberg & Bremer (2003) found the same, adding a morphological matrix, but again with only weak support. *Rousseaceae* on the other hand were shown by Soltis & al. (2000) to be the sister group to *Campanulaceae,* but appear to be more distant in other broad-scale analyses (Savolainen & al. 2000; Albach & al. 2001; Bremer & al. 2002). Sampling a greater number of genomic regions, Winkworth & al. (2008) now provide increased evidence for *Campanulaceae* being sister to *Rousseaceae,* whereas *Pentaphragma* is sister to the remainder of *Asterales*. Sequence data of *petD* (this study) also depict *Pentaphragma* as distant from *Campanulaceae*.

The pre-cladistic assumption of close relations between *Sphenoclea* Gaertn. and *Campanulaceae* (e.g., Cronquist 1988) are refuted by all molecular phylogenetic analyses. In our *petD* trees that were rooted with *Solanaceae* (*Atropa* L., *Nicotiana* L.), *Sphenoclea* is shown as sister to the *Asterales* (Fig. 4, 5). Although our sampling beyond *Campanulaceae* is restricted and does not permit further insights into the exact position of *Sphenoclea,* it provides clear arguments against any close association with *Campanulaceae*.

This phylogenetic analysis of the bluebell family has a representative taxon sampling that includes *Cyphioideae* together with *Lobelioideae* and a broad sampling of *Campanuloideae*. Previous phylogenetic analyses either rooted *Campanulaceae* with *Lobeliaceae* (e.g., Eddie & al. 2003), just focused on *Campanula* and putative close relatives using *Jasione* L., *Roella* L. and *Wahlenbergia* Roth as outgroups (Roquet & al. 2008), or sampled *Lobelioideae* and very few *Campanuloideae* with other *Asterales* as outgroups (e.g., Antonelli 2008). Using molecular (*atpB, ndhF, rbcL*) and phenotypic characters to analyse phylogenetic relationships of *Asterales,* Lundberg & Bremer (2003) included one sample of *Campanula* and *Codonopsis* Wall. to represent *Campanuloideae,* and *Cyphia rogersii* S. Moore, along with three genera of the *Lobelioideae*. Whereas *Cyphia* P. J. Bergius appeared either unresolved or as sister to *Campanula* plus *Codonopsis,* the monophyly of *Campanulaceae* s.l. was clearly established. Our Maximum Parsimony reconstruction of *petD* sequence data finds *Cyphioideae* as weakly supported sister to *Campanuloideae* (Fig. 4), whereas Bayesian Inference is inconclusive (Fig. 5). The other two small subfamilies *Cyphocarpoideae* and *Nemacladoideae* (Lammers 2007) have not been included so far in molecular analyses. APG II (2003) does not exclude the possibility of keeping *Campanulaceae* and *Lobeliaceae* as separate taxa, but it appears that the problem has not yet been thoroughly investigated. The position of the other three subfamilies in the *Campanulaceae* clade needs to be clarified before any decision on internal classification can be made.

Phylogenetic relationships within the core *Campanuloideae* and *Campanula*

Core *Campanuloideae* — Our study depicts a well supported clade of core *Campanuloideae* constituted by four major lineages (Fig. 4, 5). However, the relationships of the *Jasione* and *Musschia* clades with respect to the two other large *Campanula* clades are not yet clear. The *Canarina* and *Wahlenbergia* clades appear as successive sisters to core *Campanuloideae* with maximum support. The ITS trees of Eddie & al. (2003) and Roquet & al. (2008), which comprised a more representative sampling of *Campanuloideae* than previously available phylogenetic analyses, did not resolve the backbone under Maximum Parsimony. Nevertheless, Eddie & al. (2003) provided evidence for a so-called *Rapunculus* clade (corresponding to our well supported sister group of the *Campanula rotundifolia* + *Phyteuma* clades) with the genera *Githopsis* Nutt. and

Fig. 5. Bayesian phylogram of *Campanulaceae* (majority rule tree found with four independent runs) based on the *petD* data set including indels. – Posterior probabilities are given above branches.

Heterocodon Nutt. as weakly supported sister group. What we call the *Campanula latifolia* clade (Fig. 4, 5) is contained for the most part in their "*Campanulaceae* s.str. clade". The *Jasione* and *Musschia* clades, including *Wahlenbergia,* appeared unresolved as "transitional taxa". Bayesian Inference by Roquet (2008) then improved the picture for the *Campanula* s.str. and *Rapunculus* clades. However, in a combined ITS and *trnL-F* tree the latter authors showed an incongruent topology, with *Wahlenbergia* branching from inside the core campanuloids (Roquet & al. 2008). Since the authors also used outgroups that at least in the case of the *Jasione* clade belong to the core *Campanuloideae,* the respective results need to be viewed with care. To summarise, both the *Jasione* and the *Musschia* clades have to be included along with representatives of the two main *Campanula* clades into the core *Campanuloideae* for further testing the phylogeny of major clades of *Campanula.*

The two major campanuloid clades are the *Campanula latifolia* + *Azorina* clade (comprising the nomenclatural type of *Campanula*) and the *C. rotundifolia* + *Phyteuma* clade. In this study, we will not further discuss the subgeneric and sectional classification of *Campanula* as this has been done by Eddie & al. (2003) and Shulkina & al. (2003), who also provide detailed overviews of the classification history of *Campanula* and so far existing morphological, anatomical and cytological studies.

Campanula lactiflora M. Bieb. is sister to *Musschia* Dumort. with high confidence (Fig. 4, 5), corroborating the ITS tree of Eddie & al. (2003). Shulkina (1979) placed this species into its own genus *Gadellia* Schulkina, endemic to the Caucasus. However, other close relatives (*C. primulifolia* Brot., *C. peregrina* L.) could not be sampled here. Segregation of these species from *Campanula* could be justified only if *Musschia* is upheld as a separate genus, and this again requires clarification of the relationships of the four major lineages in core *Campanuloideae.*

Campanula latifolia clade — *Campanula rumeliana* (Hampe) Vatke (≡ *C. jacquinii* subsp. *rumeliana* (Hampe) Kit Tan) is resolved as isolated sister lineage (94 % JK, 1.0 PP) to all other members of this clade (90 % JK, 1.0 PP; Fig. 4, 5). Our topology is congruent with the nrITS tree of Roquet & al. (2008) and the 3-gene analysis of Cellinese & al. (2009) but *petD* yields better node support (nrITS: no statistical confidence; *rbcL*+*atpB*+*matK*: 80 % JK, 1.0 PP and 81 % JK, 1.0 PP, respectively). *C. rumeliana* has a capitate inflorescence with very narrow-infundibuliform corolla (less than 2 mm broad) and was therefore described and traditionally filed under *Trachelium* L. (Tan & Iatrou 2001) or under its segregate *Diosphaera* Buser (Hayek 1928-31). It is endemic to S Bulgaria and NE Greece. Other molecular studies found *Trachelium*

caeruleum L. (the nomenclatural type of *Trachelium*) as sister to the *C. latifolia* + *Azorina* clade (Roquet & al. 2008; Cellinese & al. 2009), underscoring the distinctness of *Diosphaera* from *Trachelium.* Thus, *Trachelium,* as treated for example in Flora Europaea (Tutin 1976), is not a natural group. When restricted to *T. caeruleum* (as in Greuter & al. 1984), it is a W Mediterranean genus.

Within the *Campanula latifolia* clade several groups of species are resolved, but their interrelationships still remain unclear. In the following we will discuss such lineages that can be phylogenetically defined. These lineages either comprise species that are also morphologically allied or species from certain geographical regions, which can be considered ro represent smaller radiations within *Campanula.*

The latter is true for the largest subclade (70 % JK, 1.0 PP, Fig. 4, 5), which, alongside with the widespread *Campanula bononiensis* L. and *C. rapunculoides* L. (see below), is mostly made up of Caucasian species (*C. alliariifolia* Willd. [extending to Anatolia], *C. armazica* Kharadze, *C. ciliata* Stev., *C. darialica* Kharadze, *C. fedorovii* Kharadze, *C. grossheimii* Kharadze, *C. hohenackeri* Fisch. & al., *C. hypopolia* Trautv., *C. kolenatiana* C. A. Mey ex Rupr., *C. makaschvilii* E. Busch and *C. petrophila* Rupr., *C. saxifraga* M. Bieb.), indicating a secondary radiation of *Campanula* in that region. Relationships between *C. alliariifolia, C. grossheimii* and *C. hohenackeri* were also depicted in the ITS trees of Eddie & al. (2003) and Roquet & al. (2008), in the latter gaining higher support with a different, improved taxon sampling. A second, smaller lineage of predominantly Caucasian species is composed by *C. annae* Kolak., *C. collina* Sims, *C. sarmatica* Ker Gawl. and the amphi-Adriatic *C. foliosa* Ten. (Fig. 4, 5).

Campanula latifolia L. and *C. trachelium* are resolved in a clade, in the Bayesian tree of our study (Fig. 5; 0.8 PP), that has not been found by other molecular phylogenetic analyses. Eddie & al. (2003) point out that these belong to a morphologically distinctive group of Eurasian mesophytic species with elongate, spicate inflorescences and lacking appendages, also including *C. bononiensis* and *C. rapunculoides.* However, the latter species belong to the aforementioned mostly Caucasian clade in our study, and also in the ITS tree of Roquet & al. (2008), as far as *C. rapunculoides* is concerned (*C. bononiensis* was not sampled; see below for the surprisingly different position of *C. trachelium* in the *trnL-F* tree of Roquet & al. 2008). *C. bononiensis* and *C. rapunculoides* are species with a European to W Siberian distribution that also occur in the Caucasus (Podlech 2008; Kolakovsky 1992).

Edraianthus (A. DC.) DC. clearly is nested within the *Campanula latifolia* clade in our trees, corroborating earlier findings (Eddie & al. 2003; Shulkina & al. 2003; Roquet & al. 2008). Phylogenetic studies thus agree on its placement within *Campanuloideae,* as sug-

gested by Fedorov (1957) and Kolkovsky (1987, 1994) and not close to *Wahlenbergia* where it was placed historically (Candolle 1830; Schönland 1889; Kovanda 1978; Takhtajan 1997). In a detailed analysis of this Balkan centred group of 10-28 species, Stefanović & al. (2008) showed the monophyly of *Edraianthus* except the Caucasian species *E. owerinianus* Rupr. (≡ *Muehlbergella oweriniana* (Rupr.) Feer; also accepted as generically distinct by Lammers 2007) and the N Greek *E. parnassicus* (Boiss. & Spruner) Halácsy (≡ *Halacsyella parnassica* (Boiss. & Spruner) Janch.), using *trnL-F* and *atpB-rbcL* spacer sequences. The apical capsule dehiscence that was considered characteristic for *Edraianthus* (e.g., by Lakušić 1974) thus is a derived feature of several species, whereas a clear phenotypic synapomorphy for the emended monophyletic treatment of the genus is still not known (Stefanović & al. 2008). Hypotheses about the next relatives of *Edraianthus* are, however, difficult. Our study depicts *Campanula barbata* as sister lineage albeit with low support (0.7 PP). Roquet & al. (2008) show *Edraianthus* in a well supported clade with *C. medium* L., *C. spicata* L., *C. incurva* A. DC. and *C. hofmannii* (Pant.) Greuter & Burdet (≡ *Symphyandra hofmanii* Pant.). The ITS trees in Eddie & al. (2003) are inconsistent, with *C. latifolia* and *Edraianthus* as sisters (no statistical support). Bayesian reconstruction on the 3-gene data set (Cellinese & al. 2009) shows *E. graminifolius* A. DC. in a grade after *C. medium* and followed by a lineage of *Symphyandra armena* A. DC. and *C. saxatilis* L. albeit again without statistical confidence. Stefanović & al. (2008) show *C. betulifolia* C. Koch, *C. radicosa* Bory & Chaub., *C. seraglio* Kit Tan & Sorger and *C. tymphaea* Hausskn. as close relatives, but the latter three have not been included in any other phylogenetic study so far. Since none of the so far published analyses includes all of the aforementioned species, a sensible comparison of phylogenetic hypotheses derived from the different data sets is difficult and may have to await a more complete species sampling of *Campanula* s.l.

A group with *Campanula spicata, C. thyrsoides* L., and *C. tridentata* Schreb. is well resolved with *petD* but *C. spicata* has a different, incongruent position in the ITS trees, appearing in the same subclade as *C. medium* and *Edraianthus* (Roquet & al. 2008). The *trnL-F* tree depicts *C. spicata* in yet another place, as sister to *C. alliariifolia* and *C. trachelium*. On the other hand, *C. spicata, C. thyrsoides* and *C. tridentata* share morphological characters such as long, more or less compact inflorescences with upright flowers and acuminate calyx teeth, supporting a hypothesis of their close relationship.

Campanula cervicaria L. and *C. glomerata* L. are found to be closely related. *C. glomerata,* a taxonomically complex species with different subspecies in various parts of Europe (Fedorov & Kovanda 1976), and *C. cervicaria* also show morphological similarities with the previous lineage. Much additional work will be nec-

essary to clarify species limits within this group and to test its relationships to *C. spicata* and its allies.

The lineage of *Campanula carpatha* Halácsy and *C. tomentosa* Vent. is also newly found in this *petD* study. *C. tomentosa* is a chasmophytic species with quinquelocular ovary, endemic to the calcareous massif of Samsun Dağ in coastal W Anatolia, between the ancient cities of Ephesus and Priene. *C. carpatha,* endemic to the SE Aegean islands of Karpathos and Saria, belongs to a group of biennial or short-lived perennial, five-carpelled chasmophytes confined to the S Aegean island arc (Phitos 1965). Its nearest relatives are *C. pelviformis* Lam. and *C. tubulosa* (Boiss.) Engl. All three are endemics of the Cretan area and are considered to form a radiation of neo-endemics (Cellinese & al. 2009).

Five other E Mediterranean species belong to the *Campanula latifolia* clade (Fig. 4, 5, see bottom part of uppermost clade), but still with unclear relationships. (1) *C. incurva* A. DC. is a perennial, large-flowered endemic of E Central Greece. Among the three-carpelled species of *Campanula* with appendiculate calyx sinuses it is unique by its cordate leaf blades and is excluded by Phitos (1964) from the Greek tricarpellary species of *C.* sect. *Medium* studied by him. (2) *C. sartorii* Boiss. & Heldr. is endemic to the island of Andros (W Kiklades). It is three-carpelled, with an exappendiculate calyx, which is the normal condition in *Campanula*. In habit and leaf shape it recalls some amphi-Adriatic members of the *C. elatines* aggregate (Greuter & al. 1984), from which it was, however, excluded by Damboldt (1965) due to differences in seed morphology. The studied members of the latter group belong in the *C. rotundifolia* clade (see below). (3) *C. thessala* Maire belongs to the Greek tricarpellary species of *C.* sect. *Medium* studied by Phitos (1964). Greuter & al. (1984) place it in the *C. rupestris* aggregate, a group of 20 closely allied, predominantly five-carpelled taxa of facultative chasmophytes with petiolate, lyrate to lobate rosette leaves and appendiculate calyx sinuses, confined to Greece and the Aegean islands, revised by Phitos (1964, 1965). (4) *C. hierapetrae* Rech. f., a local endemic of Mt Afendis Kavousi in E Crete, belongs to the *C. heterophylla* aggregate (Greuter & al. 1984): four rare and local Aegean island endemics, to which the recently described *C. koyuncui* H. Duman of Baba Dağ in SW Anatolia is to be added (Fielding & al. 2005). Cellinese & al. (2009) found *Michauxia tchihatchewii* Fisch. & Heldr. to be sister to *C. hierapetrae*. (5) In our *petD* tree, *M. campanuloides* L'Hér. is also unresolved in the *C. latifolia* clade. *Michauxia* L'Hér. is a genus of seven species growing in SW Asia, predominantly Turkey (Damboldt 1978; Lammers 2007). Its main characteristic, the conspicuous stigma with 8-10 recurved lobes, appears to be a derived feature, and likely relationships to E Mediterranean *Campanula* taxa have to be further tested. Surprisingly, ITS data resolve *M. tchihatchewii* sister to *C. barbata* (Alps and Pyrenees,

Fig. 1A), with variable support (Eddie & al. 2003, 98 % BS; Roquet & al. 2008, 0.79 PP). Future studies will have to evaluate whether there might be incongruence between plastid and nuclear DNA trees that would point to ancient reticulations.

***Azorina* clade** — Our study shows a well defined *Azorina* clade, indicating close relationships of the Azorean endemic genus *Azorina* Feer, Cape Verdean *Campanula* species (*C. bravensis* Bolle, *C. jacobaea* C. Sm.; both included for the first time in a molecular phylogenetic study) and an E African species (*C. edulis* Forssk.). This clade was anticipated by Eddie & al. (2003), because all these species belong to a group with basal capsule dehiscence, appendages between the calyx lobes (e.g., Thulin 1976), and a total area ranging from the mid Atlantic archipelagos via the E Mediterranean to E Africa. The clade is not readily comparable with results from Roquet & al. (2008) due to a strongly deviating taxon sampling, but their ITS and *trnL-F* trees show the Ibero-Maghrebine *C. mollis* L. (a relative of *C. edulis*) as a close relative of *Azorina*.

***Campanula rotundifolia* clade** – This well supported clade as understood in our study (Fig. 4, 5) includes most species with bell-shaped flowers and is sister to the *Phyteuma* clade, which contains many species with a corolla deeply split into narrow segments or with a closed tubular corolla (see below). The *C. rotundifolia* aggregate (Greuter & al. 1984; Fedorov & Kovanda 1976; Fig. 1C) forms a well supported clade including species such as *C. rotundifolia* (widespread), *C. scheuchzeri* Vill. (Pyrenees, Alps, W Carpathians, Balkans) and *C. baumgartenii* J. Becker (W Central Europe; Buttler 2002). Relationships within this clade are difficult to resolve, since it contains many polyploids (Kovanda 1970a, b, 1977) and species limits are unclear in many cases. Further work will thus have to evaluate possible reticulate patterns and incomplete lineage sorting, using a geographically dense sampling. The dwarf mountain species, represented by *C. cespitosa* Scop. and *C. cochleariifolia* Lam., appear as close but independent relatives (Fig. 4, 5) of the phylogenetically defined *C. rotundifolia* alliance. ITS data did not provide resolution for these species (Park & al. 2006). The North American *C. divaricata* Michx. is a third member of this group, corroborating earlier findings of Eddie & al. (2003) based on ITS and indicating a relatively recent migration to North America from Eurasia. The *C. elatines* aggregate (Greuter & al. 1984), comprising the isophyllous species studied by Damboldt (1965), is polyphyletic in the *petD* trees of our study, in agreement with the nuclear ITS trees (Park & al. 2006): *C. elatinoides* Moretti, *C. fragilis* Cyr. and *C. isophylla* Moretti are resolved as relatives to the *C. rotundifolia* alliance, whereas a second clade with *C. elatines* L. and *C. portenschlagiana* is more distant.

The latter species was found within the highly supported *C. garganica* clade by Park & al. (2006).

A novel result of our study is the well supported sister-group relationship of *Campanula patula* L. and *C. spatulata* Sm. Previous phylogenetic analyses never included both species together. *C. patula*, predominantly biennial but with a perennial subspecies to the south (Carpathians and N Balkans), is widespread in European woodlands, whereas *C. spatulata,* a perennial geophyte with a napiform taproot, is endemic to the S Balkan Peninsula (two altitudinally vicarious subspecies) and Crete (a third subspecies). The surface parts of these two vicarious species are similar. Cano-Maqueda & al. (2008) found *C. patula* and *C. rapunculus* to be sisters, with the two in turn being the sister group to the W Mediterranean monophyletic *C. lusitanica* alliance. The whole clade (including *C. patula* and *C. rapunculus*) was called *C. lusitanica* lineage by Cano-Maqueda & al. (2008), distant from the *C. dieckii* lineage that is geographically restricted to the Baetic range. More extensive taxon sampling is needed to assess the relationship of the E Mediterranean *C. spatulata* to the *C. lusitanica* lineage, suggested by the position of *C. patula* in our study (Fig. 4, 5). Roquet & al. (2007) also found that W Mediterranean taxa of *C.* sect. *Rapunculus* often have a sister-group relationship with E Mediterranean and W Asian taxa.

A clade containing *Campanula carpatica* Jacq., *C. pyramidalis* L. and *C. versicolor* is of note. All have large flowers and the latter two are sister species and share very open flowers (Fig. 1D). *C. versicolor* is a highly variable, sturdy perennial with a woody rootstock, mostly growing as a chasmophyte from NE Italy to the Balkan Peninsula and the Ionian Islands. *C. carpatica* (Carpathian mountains) and *C. pyramidalis* (N Italy and NW Balkan Peninsula) grow in similar habitats. The lineage thus appears to reflect a SW European to E Mediterranean species radiation. *C. carpatica* and *C. pyramidalis* had been found without support as sister groups with ITS (Eddie & al. 2003; Roquet & al. 2008), and with high support in the 3-gene analysis (Cellinese & al. 2009). *C. versicolor* had not been previously studied, but was considered as a member of the *C. pyramidalis* aggregate by, e.g., Greuter & al. (1984). A putative fossil relative of *C. pyramidalis* and *C. carpatica,* called *C. palaeopyramidalis* (Łańcucka-Środoniowa 1977, 1979), the only so far known *Campanula* fossil, can be used for internally calibrating the tree (Cellinese & al. 2009).

***Phyteuma* clade** — The *Campanula persicifolia* lineage occupies an isolated position as sister to all remaining members of the *Phyteuma* clade in this study (Fig. 4, 5, maximum support) and all other published molecular studies (variable support). Eddie & al. (2003) and Roquet & al. (2008) furthermore show that *C. stevenii* M. Bieb. (E Europe) and *C. pterocaula* Hausskn. (Tur-

key) are part of the crown group for this lineage. *C. persicifolia* L. is a large-flowered species with a wide European range. Branching after *C. persicifolia* in the *Phyteuma* clade is *C. trichocalycina* Ten. It is a perennial peculiar in having a bluish violet corolla deeply divided into linear lobes, not unlike those of *Petromarula* R. Hedw. or *Asyneuma* Griseb. & Schenk. Its taxonomic position remains controversial: first described under *Campanula,* the species has variously been assigned to *Podanthum* (G. Don) Boiss.*, Phyteuma* L. or *Asyneuma* as recently again adopted by Conti & al. (2005). Lakušić & Conti (2004) subdivided it into *A. trichocalycinum* s.str. (S Italy and Sicily) and *A. pichleri* (Vis.) D. La - kušić & F. Conti (Balkan Peninsula and Crete), to which latter the material we studied pertains. *Petromarula pinnata* (Fig. 1E) is a monotypic genus endemic to Crete and is rather isolated morphologically, being unique among European *Campanulaceae* in having compound leaves. ITS data (Roquet & al. 2008) and the 3-gene analysis (Cellinese & al. 2009) resolve it as sister to the *Physoplexis-Phyteuma* clade, a position which, for once, is not fully corroborated by the *petD* trees (Fig. 4, 5). The astonishing position of *Legousia* Durande as sister to *Petromarula* (Fig. 4) is not corroborated by the Bayesian phylogram (Fig. 5) and should be viewed with care. *Phyteuma* appears monophyletic, bearing in mind that only four of the 24 European species (Damboldt 1976) were sampled, and appears sister to the monotypic *Physoplexis* (Endl.) Schur, endemic to the SE Alps.

Phylogenetic utility of *petD* sequences in *Campanula*

The *petD* marker was first characterized in a study of early branching angiosperms (Löhne & Borsch 2005) and has since been applied successfully to infer deep nodes in eudicots (Worberg & al. 2007) and relationships in the orders *Nymphaeales* (Löhne & al. 2007) and *Malpighiales* (Korotkova & al. 2009). Notably, recent analyses in *Rubiaceae,* tribe *Spermacoceae,* revealed *petD* sequences to contain much better phylogenetic signal than other non-coding regions (such as *trnL-F* or *rps*16) (Kårehed & al. 2008; Groeninckx & al. 2009). This is the first study using a larger set of *petD* sequences at the species level.

Although the comparison of phylogenetic structure in the *petD* data set with other data sets for *Campanulaceae* is currently difficult because *R* is strongly influenced by taxon sampling and published molecular data sets deviate considerably from our data set, a trend appears to be obvious. *R* is a measure for the average statistical support related to the total number of nodes in a tree that can be expected from a taxon set when the tree is fully resolved (Müller & al. 2006). The maximum parsimony tree based on ITS as shown by Eddie & al. (2003) depicts several lineages in *Campanuloideae,* such as the *Platycodon* clade (100 % JK), the sister-group relationship of the *Campanula rotundifolia* + *Phyteuma* clade (100 %), the *Jasione* clade (100 %) and

the *C. latifolia* clade, each of which is hold together by its firmly similar sequence block. However, deep nodes describing relationships between these clades are not resolved with confidence. Within each clade, only few sublineages are supported. ITS trees in *Campanulaceae* are not very well resolved, and an examination of the sequences available in GenBank indicates that unlike plastid spacers and introns large parts of the ITS1 and ITS2 spacers cannot be aligned reliably between major clades. Cano-Maqueda & al. (2008) also hint to a lower performance of ITS compared to *trnT-L* sequence data. To the contrary, most nodes of the backbone of *Campanulaceae* in trees derived from *petD* sequence data receive high JK values.

Equal or better phylogenetic performance of *petD* alone is apparent when compared to trees inferred from combined genes (*rbcL+atpB+matK*) such as the 3-gene data set by Cellinese & al. (2009). Compared to the three genes *petD* has < 20 % of sequenced nucleotides. Nevertheless, resolution and support of the respective Bayesian reconstructions (Fig. 5 in this study; Cellinese & al. 2009: fig. 3) is largely similar in the overall topologies for the *Campanuloideae*. There are even some nodes with higher support in the *petD* tree, such as the crown group node for the *Campanula rotundifolia* clade (0.81 PP in the 3-gene analysis; 1.0 PP in the *petD* study), the crown group node for the *Phyteuma* clade (0.83 PP versus 1.0 PP) or the node depicting the sister-group relationship of the two aforementioned clades (0.81 PP versus 1.0 PP). The results obtained here indicate a similar performance as studies on basal angiosperms based on non-coding regions such as *trnT-trnF* or *petD* (Borsch & al. 2003; Löhne & Borsch 2005), which yielded phylogenetic hypotheses comparable to those based on five to nine genes, with a four to seven times higher amount of sequence data applied (Qiu & al. 1999, 2005). The practical aspects of "phylogenetic utility" such as easy amplification and sequencing will be addressed below.

Relative mutational rates and resolution contrast between major lineages of *Campanula*

The rate of molecular evolution in a genomic region (both substitutions and microstructural mutations) can vary over time within a lineage and mutational rates can shift with the evolution of new lineages, leading to the well known "rate heterogeneity". As a first approximation we calculated average relative substitutional rates in the *petD* intron data set for all major *Campanulaceae* groups found in this study based on a maximum likelihood estimate (GTR+Γ+I model) of substitutions per site between taxa. Significantly lower relative substitution rates were found in the *Campanula latifolia* clade (Fig. 6), which also shows lower internal branch lengths. The causes for mutational rates are different. In addition to a widespread trend of short-lived plants to exhibit higher rates (e.g., Smith & Donoghue 2008; con-
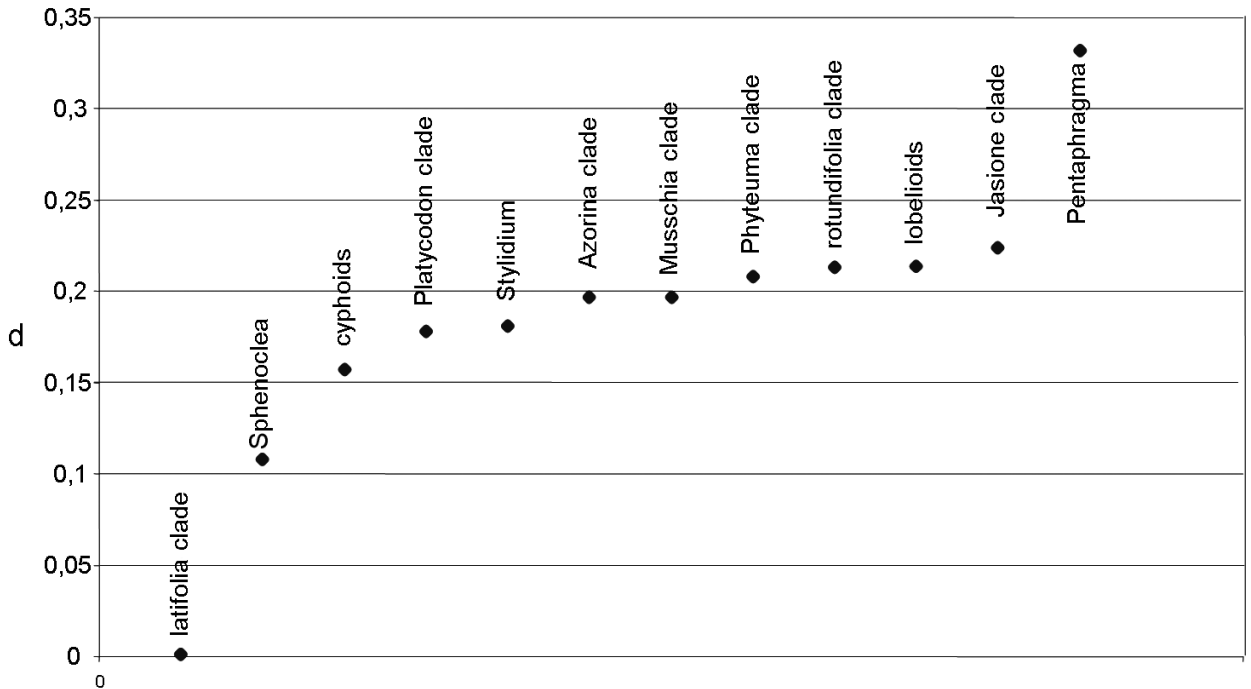
Fig. 6. Rate plot comparing relative molecular substitutional rates across major clades in *Campanulaceae* and allies with *Nicotiana tabacum* as reference taxon and *Atropa bella-donna* as outgroup. – K(ij) = maximum likelihood estimate (GTR+Γ+I model) of substitutions per site between taxa.

firming the "generation-time hypothesis"), other factors influencing the underlying patterns of fixation of mutations such as differences in DNA repair efficiency, fidelity of polymerases and metabolic changes leading to altered selective pressures are discussed.

Mutational rates within genomic regions vary by a factor of six to eight throughout angiosperms as has been shown for example for the *matK* gene (Müller & al. 2004). Molecular mutational rates also influence the performance of phylogenetic markers, through different amounts of phylogenetic structure. When average mutational rates of a genomic locus in a radiating clade are high, this will lead to a higher probability for mutations to accumulate along internal branches of this radiation as compared to when mutational rates are low. Lineage specific resolution contrasts have been found in *Piperaceae* (good resolution in *Peperomia* versus poor resolution in *Piper;* Wanke & al. 2007) and in *Lentibulariaceae* (good resolution in the *Genlisea-Utricularia* clade versus poor resolution in *Pinguicula;* Jobson & al. 2003; Müller & al. 2004). In both cases high mutational rates in the genome correlate to high resolution and support in the respective tree compartments. In a similar way, rate heterogeneity seems to lead to a resolution contrast in *Campanulaceae*. The poorly resolved *Campanula latifolia* clade also exhibits significantly lower molecular evolutionary rates in *petD* (Fig. 6).

Cellinese & al. (2009) found that a molecular clock in each of the partitions of their 3-gene data set (*rbcL+atpB+matK*) had to be rejected based on likelihood ratio tests. However, using a relaxed molecular clock in a Bayesian environment (with the program BEAST), the authors found much younger divergence times of the taxa within the *Campanula latifolia* clade (major splits around 10-15 Ma) as compared to the *C. rotundifolia* and *Phyteuma* clades (major splits around 20-40 Ma). The respective effects of time differences in diversification (i.e., the speed of organismic radiation) versus changes in underlying molecular evolutionary rates throughout the evolutionary history of *Campanula* need to be carefully evaluated. It will therefore be relevant to all dating analyses and to the analysis of species diversification how effectively shifts in molecular mutational rates can be corrected by means of approaches such as "rate smoothing" through penalized likelihood (Sanderson & al. 2002), or the application of "relaxed clocks" in a Bayesian framework (see Renner 2005 for review). Temporal shifts in species diversification rates may also be tested in a maximum likelihood framework (e.g., Rabosky 2006). Future work will, in particular, have to test the respective effects of rapid species diversification and lowered mutational rates within the *C. latifolia* clade.

Utility of *petD* for molecular species identification in *Campanula* s.l.

Sequences of *petD* (the group II intron and the *petB-petD* spacer) allow to recognize about 90 % of the species of *Campanuloideae* included in this study. Moreover, the marker is of practical utility as barcode (see, e.g., the recent discussion by Devey & al. 2009 and

Borsch & Quandt 2009). *PetD* is a genomic region that proved to be easily amplified in *Campanulaceae* using the universal primers of Löhne & Borsch (2005), also from herbarium specimens (up to 20 years old). Fragments of about 900 nt in size could be easily sequenced. The absence of larger polyA/T stretches is relevant for getting long and unambiguous reads. Using an ABI3730XL automated sequencer, reads of >800 nt could be obtained without any N, covering both strands with one primer respectively. For barcoding application, shorter areas could also be selected. Watts & al. (2008) designed universal primers for amplification of the D4 loop of chloroplast introns. A similar approach was carried out by Taberlet & al. (2007) for the P8 loop of the *trnL* intron. However, applicability depends on the level of taxonomic resolution needed for a particular question. Moreover, further work will be necessary to test performance of shorter *petD* fragments at the species level.

Contrary to phylogenetic analysis, sequence elements from mutational hotspots can be included for species identification. One of the most important issues for accurate species identification with DNA sequences, however, is to evaluate intraspecific variability, for example through assessing haplotype variation in an approach integrating phenotypic and molecular characters. The present study gives several examples (e.g. the *Campanula rotundifolia* alliance s.str., the *C. patula* + *C. spatulata* clade, the *C. cervicaria* + *C. glomerata* clade, the *C. spicata* + *C. thyrsoides* clade) for lineages where species might have diversified faster than *petD* sequences but where, essentially, species limits are still unclear. Work for the coming years therefore needs to first integrate molecular and morphological approaches for achieving robust monographic information on the species in *Campanula*.

5. Conclusions and future work

Sequence data of the *petD* group II intron allowed to generate well resolved trees at the level of species and closely related genera within the *Campanulaceae,* further underscoring the utility of this marker as a phylogenetic tool. For several nodes of the tree of *Campanula* the gene tree of this non-coding region was better supported than nodes inferred by a much larger multi-gene data set. Our study therefore provides yet another example of the high potential of non-coding and rapidly evolving chloroplast DNA for resolving phylogenetic relationships. Nevertheless, the phylogenetic structure needs to be compared for different chloroplast genomic regions, especially introns and spacers, using identical taxon sets. Such comparative analyses of molecular evolution will allow us to get further insights into structural and mutational patterns that influence the accumulation of historical signal. Group II introns with a high proportion of stem-loops (i.e., the longer introns such as *rpl*16) appear to be particularly promising. Considering the re-

sults of this and other recent studies it will certainly be possible to select genomic regions as markers for high performance. In line with a more detailed argumentation by Borsch & Quandt (2009) a good option will be to combine high performing intron and spacer data sets. A combination of several highly performing markers will certainly be needed to generate well resolved species level phylogenies for diverse genera. Fully resolved trees are, however, needed from both the organelle and nuclear genomic compartments to detect possible incongruence as evidence for reticulations in genome evolution. Hybridization and introgression may have played an important role in *Campanula,* as known series of polyploidy in some lineages and also extant hybrids indicate.

Campanula s.l. certainly is one of the most extreme cases in terms of a genus that is paraphyletic to a number of currently accepted other genera, with proposals to re-classify having been made in either the splitting or lumping direction (see Roquet & al. 2008). Results from phylogenetic analysis imply a highly adaptive evolution of floral morphology, rapid acquisition of features that adapt species to a habitat, both radiations of species groups in geographically confined areas and long distance migrations of ancestors leading to geographically remote ranges of extant, closely related species. As a consequence, there is little predictability in pre-cladistic classification systems to guide the design of taxon sampling, such as providing an answer to the question of which species to include for studying a given question. In addition, there are widespread, variable species that appear to be closely related to ecologically specialised endemics in certain regions.

Research in several directions appears necessary. First of all, relationships within core *Campanuloideae,* among the two major radiations of *Campanula* and the *Jasione* and *Musschia* clades, need to be clarified. This is the prerequisite for any sound revision of the classification system. Second, the taxon sampling has to be considerably increased to assess, whether there are any further isolated lineages in core *Campanuloideae* and to further evaluate the composition of their major clades. Many traditionally recognized species groups should be defined phylogenetically before examining the often unclear species limits. A comprehensive effort based on one or two genomic regions that are chosen for their high phylogenetic utility now appears to be feasible, with sampling extended to all the 300-400 species, in order to generate a phylogenetic framework for future evolutionary and monographic research in *Campanula*.

References

Albach D. C., Soltis P. S., Soltis D. E. & Olmstead R.G. 2001: Phylogenetic analysis of asterids based on sequences of four genes. – Ann. Missouri Bot. Gard. **88:** 163-212. [CrossRef]

Antonelli A. 2008: Higher level phylogeny and evolutionary trends in *Campanulaceae* subfam. *Lobelioideae:* molecular signal overshadows morphology. – Mol. Phylogenet. Evol. **46:** 1-18. [CrossRef]

APG II 2003: An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants. – Bot. J. Linn. Soc. **141:** 399-436. [CrossRef]

Barfuss M. H. J., Samuel R., Till W. & Stuessy T. F 2002: Phylogenetic relationships in subfamily *Tillandsioideae (Bromeliaceae)* based on DNA sequence data from seven plastid regions. – Amer. J. Bot. **92:** 337-351. [CrossRef]

Barfuss M. H. J., Samuel R., Till W. & Stuessy T. F 2005: Phylogenetic relationships in subfamily *Tillandsioideae (Bromeliaceae)* based on DNA sequence data from seven plastid regions. – Amer. J. Bot. **92:** 337-351. [CrossRef]

Belyaev A. A. 1984a: Anatomiya semyan nekotorykh predstanitelei semeistva *Campanulaceae* [Seed anatomy in some representatives of the family *Campanulaceae*]. – Bot. Zhurn. (Moscow & Leningrad) **69:** 585-594.

Belyaev A. A. 1984b: Ul'trastruktura poverkhnosti i nekotorye morphologicheskie kharakteristiki semyan predstanitelei semeistva *Campanulaceae* [Surface ultrastructure and some morphological characteristics of seeds in the representatives of the family *Campanulaceae*]. – Bot. Zhurn. (Moscow & Leningrad) **69:** 890-898.

Borsch T. & Quandt D. 2009: Mutational dynamics and phylogenetic utility of non-coding chloroplast DNA. – Pl. Syst. Evol. [CrossRef]

Borsch T., Hilu K. W., Quandt D., Wilde V., Neinhuis C. & Barthlott W. 2003: Non-coding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. – J. Evol. Biol. **16:** 558-576. [CrossRef]

Bremer B., Bremer K., Heidari N., Erixon P., Olmstead R. G., Anderberg A. A., Kallersjo M. & Barkhordarian E. 2002: Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. – Mol. Phylogenet. Evol. **24:** 274-301. [CrossRef]

Buttler K. P. 2002: Beitrag zu Kenntnis von *Campanula baumgartenii*. – Bot. Natursch. Hessen **14:** 77-90.

Candolle A. L. P. P. de 1830: Monographie des Campanulees. – Paris.

Cano-Maqueda J., Talavera S., Arista M. & Catalán P. 2008: Speciation and biogeographical history of the *Campanula lusitanica* complex *(Campanulaceae)* in the Western Mediterranean region. – Taxon **57:** 1252-1266.

Carlström A. 1986: A revision of the *Campanula drabifolia* complex *(Campanulaceae)*. – Willdenowia **15:** 375-387.

Cellinese N., Smith S. A., Edwards E. J., Kim S.-T., Haberle R. C., Avramakis M. & Donoghue M. J. 2009: Historical biogeography of the endemic *Campanulaceae* of Crete. – J. Biogeogr. [CrossRef]

Conti F., Abbate G., Alessandrini A. & Blasi C. (ed.) 2005: An annotated checklist of the Italian vascular flora. – Roma.

Cronquist A. 1988: The evolution and classification of flowering plants. – New York.

Damboldt J. 1965: Zytotaxonomische Revision der isophyllen *Campanulae* in Europa. – Bot. Jahrb. Syst. **84:** 302-358.

Damboldt, J. 1976: *Phyteuma* L. – Pp. 95-98 in: Tutin T. G., Heywood V. H., Burges N. A., Moore D. M., Valentine D. H., Walters S. M. & Webb D. A. (ed.), Flora europaea **4.** – Cambridge, etc.

Damboldt J. 1978: *Campanula* L. – Pp. 2-64 in: Davis P. H. (ed.), Flora of Turkey and the East Aegean Islands. – Edinburgh.

Devey D. S., Chase M. W. & Clarkson J. J. 2009: A stuttering start to plant DNA barcoding: microsatellites present a previously overlooked problem in noncoding plastid regions. – Taxon **58:** 7-15.

Dunbar A. 1975a: On pollen of *Campanulaceae* and related families with special reference to surface ultrastructure. I. *Campanulaceae* subfam. *Campanuloideae.* – Bot. Not. **128:** 73-101.

Dunbar A. 1975b: On pollen of *Campanulaceae* and related families with special reference to surface ultrastructure. I. *Campanulaceae* subfam. *Cyphioideae* and subfam. *Lobelioideae; Goodeniaceae; Sphenocleaceae.* – Bot. Not. **128:** 102-118.

Eddie W. M. M., Shulkina T., Gaskin J., Haberle R. C. & Jansen R. K. 2003: Phylogeny of *Campanulaceae* s.str. inferred from ITS sequences of nuclear ribosomal DNA. – Ann. Missouri Bot. Gard. **90:** 554-575. [CrossRef]

Ehrendorfer F., Manen J.-F. & Natali A. 1994: cpDNA intergene sequences corroborate restriction site data for reconstructing *Rubiaceae* phylogeny. – Pl. Syst. Evol. **190:** 245-248. [CrossRef]

Fedorov A. A. 1957: *Campanulaceae.* – Pp. 459-475 in: Shishkin B. K. (ed.), Flora SSSR **24.** – Moscow.

Fedorov A. A. & Kovanda M. 1976: *Campanula.* – Pp. 74-93 in: Tutin T. G., Heywood V. H., Burges N. A., Moore D. M., Valentine D. H., Walters S. M. & Webb D. A. (ed.), Flora europaea **4.** – Cambridge, etc.

Fielding J., Turland N. & Mathew B. 2005: Flowers of Crete. – Kew.

Gagnidze R. 2005: Vascular plants of Georgia. A nomenclatural checklist. – Tbilisi.

Greuter W., Burdet H. M. & Long G. 1984: Med-Checklist **1.** – Genève & Berlin.

Groeninckx I., Block P. de, Rakotonasolo F., Smets E. & Dessein S. 2009: Rediscovery of Malagasy *Lathraeocarpa* allows determination of its taxonomic position within *Rubiaceae*. – Taxon **58:** 209-226.

Hamilton M. B. 1999: Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. – Molec. Ecol. **8:** 521-523.

Hayek A. 1928-31: Prodromus florae peninsulae balcanicae **2** [pp. 1-96 (1928), pp. 97-336 (1929), pp. 337-576 (1930), pp. 577-1152 (1931)]. – Repert. Spec. Nov. Regni Veg. Beih. **30(2).**

Huelsenbeck J. P. & Ronquist F. 2001: MRBAYES: Bayesian inference of phylogenetic trees. – Bioinformatics **17:** 754-755. [CrossRef]

Jobson R. W., Playford J., Cameron K. M. & Albert V. A. 2003: Molecular phylogenetics of *Lentibulariaceae* inferred from plastid *rps*16 intron and *trnL-F* DNA sequences: implications for character evolution and biogeography. – Syst. Bot. **28:** 157-171.

Kårehed J., Groeninckx I., Dessein S., Motley T. J. & Bremer B. 2008: The phylogenetic utility of chloroplast and nuclear DNA markers and the phylogeny of *Rubiaceae* tribe *Spermacoceae*. – Molec. Phylogenet. Evol. **49:** 843-866. [CrossRef]

Kelchner S. A. 2002: Group II introns as phylogenetic tools: structure, function, and evolutionary constraints. – Amer. J. Bot. **89:** 1651-1669. [CrossRef]

Kocyan A., Zhang L.-B., Schaefer H. & Renner S. S. 2007: A multi-locus chloroplast phylogeny for the *Cucurbitaceae* and its implications for character evolution and classification. – Mol. Phylogenet. Evol. **44:** 553-577. [CrossRef]

Kolakovsky A. A. 1987: Systema semeistva *Campanulaceae* starogo sveta [System of the *Campanulaceae* family from the Old World]. – Bot. Zhurn. (Moscow & Leningrad) **72:** 1572-1579.

Kolakovsky A. A. 1992: *Campanulaceae.* – Pp. 170-171 in: Meusel H. & Jäger E. J. (ed.), Vergleichende Chorologie der zentraleuropäischen Flora **3.** – Jena, etc.

Kolakovsky A. A. 1994: Konspekt sistemy semeistva *Campanulaceae* starogo sveta [The conspectus of the system of Old World *Campanulaceae*]. – Bot. Zhurn. (Moscow & Leningrad) **79(1):** 109-124.

Korotkova N., Schneider J. V., Quandt D., Worberg A., Zizka G. & Borsch T. 2009: Phylogeny of the eudicot order *Malpighiales:* analysis of a recalcitrant clade with sequences of the *pet*D group II intron. – Pl. Syst. Evol. [CrossRef]

Kovanda M. 1970a: Polyploidy and variation in the *Campanula rotundifolia* complex. Part 1 (General). – Rozpr. Československ. Akad. Ved. **80:** 1-95.

Kovanda M. 1970b: Polyploidy and variation in the *Campanula rotundifolia* complex. Part 2 (Taxonomy). 1. Revision of the groups *Saxicolae, Lanceolatae* and *Alpicolae* in Czechoslovakia and adjacent regions. – Folia Geobot. Phytotax. **5:** 171-208.

Kovanda M. 1977: Polyploidy and variation in the *Campanula rotundifolia* complex. Part 2 (Taxonomy). 1. Revision of the groups *Vulgares* and *Scheuchzerianae* in Czechoslovakia and adjacent regions. – Folia Geobot. Phytotax. **12:** 23-89.

Kovačić S. 2004: The genus *Campanula* L. *(Campanulaceae)* in Croatia, circum-Adriatic and West Balkan region. – Acta Bot. Croat. **63:** 171-202.

Lakušić D. 1974: Prirodni system populacija I vrsta roda *Edraianthus* DC. – God. Biol. Inst. Univ. Sarajevu **26:** 1-30.

Lakušić D. & Conti F. 2004: *Asyneuma pichleri (Campanulaceae),* a neglected species of the Balkan Peninsula. – Pl. Syst. Evol. **247:** 23-36.

Lammers T. G. 2007: *Campanulaceae.* – Pp. 36-56 in: Kadereit J. W. & Jeffrey (volume ed.), Families and genera of vascular plants **8.** – Berlin, etc.

Łańcucka-Środoniowa M. 1977: New herbs described from the Tertiary of Poland. – Acta Palaeobot. **18:** 37-42.

Łańcucka-Srodoniowa M. 1979: Macroscopic plant remains from the freshwater Miocene of the Nowy Sacz basin (West Carpathians, Poland). – Acta Palaeobot. **20:** 74-75.

Lehmann K. & Schmidt U. 2003: Group II introns: structure and catalytic versatility of large natural ribozymes. – Crit. Rev. Biochem. Molec. Biol. **38:** 249-303. [CrossRef]

Löhne C. & Borsch T. 2005: Molecular evolution and phylogenetic utility of the *petD* group II intron: a case study in basal angiosperms. – Molec. Biol. Evol. **22:** 317-332. [CrossRef]

Löhne C., Borsch T. & Wiersema J. H. 2007: Phylogenetic analysis of *Nymphaeales* using fast-evolving and non-coding chloroplast markers. – Bot. J. Linn. Soc. **154:** 141-163. [CrossRef]

Lundberg J. & Bremer K. 2003: A phylogenetic study of the order *Asterales* using one morphological and three molecular data sets. – Int. J. Pl. Sci. **164:** 553-578. [CrossRef]

Manen J. F., Natali A. & Ehrendorfer F. 1994a: Phylogeny of *Rubiaceae-Rubieae* inferred from the sequence of a cpDNA intergene region. – Pl. Syst. Evol. **190:** 195-211. [CrossRef]

Mathews D. H., Schröder J. S., Turner D. H. & Zuker M. 2006: Predicting RNA secondary structure. – Pp. 631-657 in: Oesteland R. F., Cech T. R. & Atkins J. F. (ed.), The RNA world, ed. 3. – Cold Spring Harbor.

Mathews D. H., Disney M. D., Childs J. L., Schroeder S. J., Zuker M. & Turner D. H. 2004: Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. – Proc. Natl. Acad. Sci. USA **101:** 7287-7292. [CrossRef]

Michel F., Umesono K. & Ozeki H. 1989: Comparative and functional anatomy of group II catalytic introns: a review. – Gene **82:** 5-30. [CrossRef]

Müller J., Müller K., Neinhuis C. & Quandt D. 2005+: PhyDE: Phylogenetic Data Editor. – www.phyde.de

Müller K. 2004a: PRAP – computation of Bremer support for large datasets. – Mol. Phylogenet. Evol. **31:** 780-782. [CrossRef]

Müller K. 2004b: GRate: relative rate test for groups of taxa using general DNA substitution models in PAUP. – http://systevol.nees.uni-bonn.de/software/GRate

Müller, K. 2005a: SeqState: Primer design and sequence statistics for phylogenetic DNA datasets. – Appl. Bioinformatics **4:** 65-69.

Müller, K. 2005b: The efficiency of different search strategies in estimating parsimony jackknife, bootstrap, and Bremer support. – BMC Evol. Biol. **5:** no. 58. [CrossRef]

Müller K., Borsch T. & Hilu K. W. 2006: Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: contrasting *mat*K, *trn*T-F and *rbc*L in basal angiosperms. – Mol. Phylogenet. Evol. **41:** 99-117. [CrossRef]

Müller K., Borsch T., Legendre L., Porembski S., Theisen I. & Barthlott W. 2004: Evolution of carnivory in *Lentibulariaceae* and the *Lamiales*. – Pl. Biol. **6:** 477-490. [CrossRef]

Park J.-M., Kovačić S., Liber Z., Eddie W. M. M. & Schneeweiss G. M. 2006: Phylogeny and biogeography of isophyllous species of *Campanula (Campanulaceae)* in the Mediterranean area. – Syst. Bot. **31:** 862-880. [CrossRef]

Phitos D. 1964: Trilokuläre *Campanula*-Arten der Ägäis. – Österr. Bot. Z. **111:** 208-230. [CrossRef]

Phitos D. 1965: Die quinquelokulären *Campanula*-Arten. – Österr. Bot. Z. **112:** 449-498. [CrossRef]

Podlech D. 1965: Revision der europäischen und nordafrikanischen Vertreter der Subsect. *Heterophylla* (Wit.) Fed. der Gattung *Campanula* L. – Feddes Repert. **71:** 50-187

Podlech D. 2008: *Campanulaceae*. – Pp. 182-322 in: Wagenitz G. (volume ed.), Gustav Hegi Illustrierte Flora von Mitteleuropa, ed. 2, **6(2A).** – Jena.

Podlech D. & Damboldt, J. 1964: Zytotaxonomische Beiträge zur Kenntnis der Campanulaceen in Europa. – Ber. Deutsch. Bot. Ges. **76:** 360-369.

Posada D. & Crandall K. A. 1998: Modeltest: testing the model of DNA substitution. – Bioinformatics **14:** 817-818. [CrossRef]

Qiu Y.-L., Lee J., Bernasconi-Quandroni F., Soltis D. E., Soltis P. S., Zanis M., Zimmer E. A., Chen Z., Savolainen V. & Chase M. W. 1999: The earlierst angiosperms: evidence from mitochondrial, plastid and nuclear genomes. – Nature **402:** 404-407. [CrossRef]

Qiu Y.-L., Dombrovska O., Lee J., Li L., Whitlock B. A., Bernasconi-Quadroni F., Rest J. S., Borsch T., Hilu K. W., Renner S., Soltis D. E., Soltis P. S., Zanis M. J., Gutell J.C., Powell M., Savolainen V., Chatrou L. W. & Chase M. W. 2005: Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. – Int. J. Pl. Sci. **166:** 815-842. [CrossRef]

Quandt D. & Stech M. 2003: Molecular systematics of bryophytes in context of land plant evolution. – Pp. 267-295 in: Sharma A. K. & Sharma A. (ed.), Plant genome: biodiversity and evolution **1.** – Enfield.

Rabosky D. L. 2006: Likelihood methods for detecting temporal shifts in diversification rates. – Evolution **60:** 1152-1164.

Renner S. S. 2005: Relaxed molecular clocks for dating historical plant dispersal events. – Trends Pl. Sci. **10:** 550-558. [CrossRef]

Rijk P. de, Wuyts J. & Wachter R. D. 2003: RnaViz2: an improved representation of RNA secondary structure. – Bioinformatics **19:** 299-300. [CrossRef]

Roquet C., Sáez L., Aldasoro J. J., Susanna A., Alcarón M. L. & Garcia-Jacas N. 2008: Natural delineation, molecular phylogeny and floral evolution in *Campanula*. – Syst. Bot. **33:** 203-217. [CrossRef]

Sanderson M. J. 2002: Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. – Molec. Biol. Evol. **19:** 101-109.

Sang T., Crawford D. & Stuessy T. F. 1997: Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia (Paeoniaceae)*. – Amer. J. Bot. **84:** 1120-1136. [CrossRef]

Savolainen V., Fay M. F., Albach D. C., Backlund M., Van der Bank M., Cameron C. M., Johnson S. A., Lledo L., Pintaud J. C., Powell M., Sheanan M. C., Soltis D. E., Soltis P. S., Weston P., Whitten W. M., Wurdack K. J. & Chase M. W. 2000: Phylogeny of the eudicots: a nearly complete familial analysis of the rbcL gene sequences. – Kew Bull. **55:** 257-309. [CrossRef]

Schönland S. 1889: *Campanulaceae*. – Pp. 40-70 in: Engler A. & Prantl K. (ed.), Die natürlichen Pflanzenfamilien **4(5).** – Leipzig.

Shaw J., Lickey E. B., Schilling E. E. & 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. – Amer. J. Bot. **94:** 275-288. [CrossRef]

Shulkina T. V. 1979: K voprosu o sistematicheskom polozhenii *Campanula lactiflora* Bieb. [De positione systematica *Campanula lactiflora* Bieb.]. – Novosti Sist. Vyssh. Rast. **16:** 175-179.

Shulkina T. V., Gaskin J. F. & Eddie W. M. M. 2003: Morphological studies toward an improved classification of *Campanulaceae* s.str. – Ann. Missouri Bot. Gard. **90:** 576-591. [CrossRef]

Simmons M. P. & Ochoterena H. 2000: Gaps as characters in sequence-based phylogenetic analyses. – Syst. Biol. **49:** 369-381. [CrossRef]

Small R. L., Ryborn J. A., Cronn R. C., Seelanan T. & Wendel J. F. 1998: The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogenetic reconstruction in a recently diverged plant group. – Amer. J. Bot. **85:** 1301-1315. [CrossRef]

Smith S. A. & Donoghue M. J. 2008: Rates of molecular evolution are linked to life history in flowering plants. – Science **322:** 86-89. [CrossRef]

Soltis D. E., Soltis P. S., Chase M. W., Mort M. E., Albach D. C., Zanis M., Savolainen V., Hahn W. H., Hoot S. B., Fay M. F., Axtell M., Swensen S. M., Prince L. M., Kress J. W., Nixon K. C. & Farris J. S. 2000: Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. – Bot. J. Linn. Soc. **133:** 381-461. [CrossRef]

Stefanović S., Lakušić D., Kuzmina M., Medeović S., Tan K. & Stefanović V. 2008: Molecular phylogeny of *Edraianthus* (Grassy Bells; *Campanulaceae*) based on non-coding plastid DNA sequences. – Taxon **57:** 452-475.

Strid A. & Tan K. (ed.) 1991: Mountain flora of Greece **2.** – Edinburgh.

Swofford D. L. 1998: PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). – Sunderland.

Taberlet P., Gielly L., Pautou G. & Bouvet J. 1991: Universal primers for amplification of 3 noncoding regions of chloroplast DNA. – Pl. Molec. Biol. **17:** 1105-1109. [CrossRef]

Taberlet P., Coissac E., Pompanon F., Gielly L., Miquel C., Valentini A., Vermat T., Corthier G., Brochmann C. & Willerslev E 2007: Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. – Nucl. Acids Res. **35:** e14 [CrossRef]

Takhtajan A. 1997: Diversity and classification of flowering plants. – New York.

Tan K. & Iatrou G. 2001: Endemic plants of Greece. The Peloponnese. – København.

Tesfaye K., Borsch T., Govers K. & Bekele E. 2007: Characterization of *Coffea* chloroplast microstellites and evidence for the recent divergence of *C. arabica* and *C. eugenioides* chloroplast genomes. – Genome **50:** 1112-1129. [CrossRef]

Thulin M. 1976: *Campanula keniensis* Thulin sp. nov., and notes on allied species. – Bot. Not. **128:** 350-356.

Timme T. E., Kuehl J. V., Boore J. L. & Jansen R. K. 2007: A comparative analysis of the *Lactuca* and *Helianthus (Asteraceae)* plastid genomes: identification of divergent regions and categorization of shared repeats. – Amer. J. Bot. **94:** 301-312.

Toor N., Hausner G. & Zimmerly S. 2001: Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. – RNA **7:** 1142-1152. [CrossRef]

Tutin T. G. 1976: *Trachelium*. – Pp. 94-95 in: Tutin T. G., Heywood V. H., Burges N. A., Moore D. M., Valentine D. H., Walters S. M. & Webb D. A. (ed.), Flora europaea **4.** – Cambridge, etc.

Wanke S., Jaramillo M. A., Borsch T., Samain M. S., Quandt D. & Neinhuis C. 2007: Evolution of the *Piperales: matK* and *trnK* intron sequence data reveals a lineage specific resolution contrast. – Molec. Phylogenet. Evol. **42:** 477-497. [CrossRef]

Watts C. D., Fisher A. E., Shrum C. D., Newbold W. L., Hansen S., Liu C. & Kelchner S. 2008: The D4 set: primers that target highly variably intron loops in plant chloroplast geomes. – Molec. Ecol. Notes **8:** 1344-1347.

Westhoff P. & Herrmann R. G. 1988: Complex RNA maturation in the chloroplast: the *psbB* operon from spinach. – Eur. J. Biochem. **171:** 551-564.
[CrossRef]

Winkworth R. C., Lundberg J. & Donoghue M. J. 2008: Toward a resolution of Campanulid phylogeny, with special reference to the placement of *Dipsacales.* – Taxon **57:** 53-65

Worberg A., Quandt D., Barniske A.-M., Löhne C., Hilu K. W. & Borsch T. 2007: Phylogeny of basal eu-

dicots: insights from non-coding and rapidly evolving DNA. – Organisms Diversity Evol. **7:** 55-77.
[CrossRef]

Worberg A., Alford M. H., Quandt D. & Borsch T. 2009: *Huerteales* sister to *Brassicales* plus *Malvales,* and newly circumscribed to include *Dipentodon, Gerrardina, Huertea, Perrottetia* and *Tapiscia*. – Taxon **58:** 468-478.

Appendix 1. Plant material used in this study. Collections made in the field first list country and locality in the case of both herbarium specimens and silica gel dried samples. Samples obtained from the living collections of the Botanic Garden Berlin-Dahlem and Bonn Botanical Gardens first list the garden accession number and then country and locality data in square brackets. Collector and collection number are given in italics, the herbarium abbreviation in parantheses. Information on the specific history of the plant individual sampled can be obtained from the label that goes with the herbarium voucher. A unique identifier is given for every DNA isolate; it follows the specimen data and precedes the EMBL/GenBank accession number.

Outgroups: *Asteraceae, Lactuca sativa* L., AP007233. *Pentaphragmataceae, Pentaphragma acuminatum* Airy Shaw, Brunei, Temburong, *Sands 5295* (B), CAM 043, FN397080. *Solanaceae, Atropa bella-donna* L., NC004561. *Nicotiana tabacum* L., NC0018792. *Sphenocleaceae, Sphenoclea zeylanica* Gaertn., BG Bonn 9160 [Ivory Coast], *Borsch 3966* (B, BONN), N516, FN397090. *Stylidiaceae, Stylidium majus* Druce, BG Bonn [Australia, S Lofty Reg., Nixon-Skinner Conservation Park Perth], *Borsch 3417* (BONN), N296, FN397091.

Campanulaceae-Campanuloideae: Adenophora triphylla (Thunb.) A. DC., BG Bonn 27427 [Japan, Hokkaido Division, Ohashi, Nayoro], *Korotkova 76* (BONN), CAM119, FN396977. *Asyneuma campanuloides* (M. Bieb. ex Sims) Bornm., Georgia, Kartli, path from Kazbegi to Devdoraki, *Korotkova 25* (B, TGM), CAM045, FN396978. *Asyneuma canescens* (Waldst. & Kit.) Griseb. & Schenk, BG Bonn 9178 [without locality data], *Korotkova 89* (BONN), CAM001, FN396979. *Azorina vidalii* Ferr., BG Bonn 830 [Azores ex BG Berlin], *Korotkova 69* (BONN), CAM019, FN396980. *Campanula alliariifolia* Willd., Georgia, Kazbegi, close to the Russian border, *Korotkova 26* (B, TGM), CAM049, FN396987. *C.* *alpina* Jacq., Slovakia, *Borsch & Košuth 3903* (B), CAM076, FN397041. *C.* aff. ***annae*** Kolak., Georgia, Kartli, village of Kobi, *Korotkova & al. 30* (B, TGM), CAM095, FN396982. *C. armazica* Kharadze, Georgia, *Schewardnadze s.n.* (BONN), CAM031, FN396988. *C. barbata* L., Italy, Lago Como, Monte S. Primo, *Borsch 3842* (B), CAM052, FN396989; Italy, Lombardia, Monte Bazena, *Borsch 3859* (B), CAM085, FN396990. *C. baumgartenii* J. Becker, BG Bonn 17230 [Germany, Rheinland Pfalz, Pfalz], *Buttler 18/1* (FR), CAM107, FN396991. *C. betulifolia* K. Koch, BG Bonn 26245 [without locality data], *Korotkova 77* (BONN),

CAM102, FN396992. *C.* aff. ***bononiensis*** L., Italy, Province Bozen, *Borsch 912* (B), CAM038, FN396983. *C. bravensis* (Bolle) A. Chev., BG Bonn 13024 [Cape Verde, Fogo, Chá das Chaldeiras, between Povoçao and Mte Velha], *Kilian & Leyens 3326* (B, BONN), CAM026, FN396993. *C.* aff. *carnica* Mert. & W. D. J. Koch, Italy, Lombardia, Mte Tremalzo, *Borsch 3847* (B), CAM054, FN396984. *C. carpatha* Halácsy, BG Berlin 173-54-07-10 [Greece, Karpathos, Vroukounda], *Gartenherbar 46201* [orig. leg. Raus & Sipman] (B), CAM125, FN396995. *C. carpatica* Jacq., Slovakia, *Borsch & Košuth 3911* (B), CAM080, FN396996. *C. cespitosa* Scop., Italy, Lombardia, Mte Bazena, *Borsch 3860* (B), CAM056, FN396994. *C. cervicaria* L., Germany, Baden-Württemberg, Hegau, Plören, *Borsch s.n.* (B), CAM037; FN396997; Italy, Lombardia, Mte Stino, *Borsch 3845* (B), CAM082, FN396999; Slovakia, *Borsch & Košuth 3905* (B), CAM078, FN396998. *C. choruhensis* Kit Tan & Sorger, BG Bonn 26246 [without locality data], *Korotkova 78* (BONN), CAM105, FN397000. *C. ciliata* Stev., Georgia, Kartli, above Bakuriani, Zchazcharo, *Korotkova 37* (B), CAM047, FN397001. *C. cochleariifolia* Lam., Germany, Bavaria, Karwendel Mts, *Borsch 3819* (B), CAM010, FN397002. *C. collina* Sims, Georgia, Kartli, *Korotkova & al. 19* (B, TGM), CAM096, FN397003. *C. darialica* Kharadze, Georgia, Kartli, path from mountain Gergeti to Kazbegi, *Korotkova & al. 23* (B), CAM098, FN397004. *C. divaricata* Michx., USA, Virginia, Blue Ridge Mts, *Borsch & Neinhuis 3283* (FR), CAM033, FN397005. *C. edulis* Forssk., Ethiopia, *Wondafrash s.n.* (B, ETH), CAM062, FN397006. *C. elatinoides* Moretti, Italy, Lombardia, Lago Iseo, *Borsch 3863* (B), CAM057, FN397008. *C. fedorovii* Kharadze, Georgia, Kartli, path from Mt Gergeti to Kazbegi, *Korotkova 48* (B), CAM046, FN397009. *C. fenestrellata* Feer, BG Bonn 25632 [Croatia, Velebit], *Korotkova 79* (BONN),

CAM111, FN397007. *C. foliosa* Ten., BG Bonn 25633 [Russia, Karbadino Balkaria], *Korotkova 80* (BONN), CAM110, FN397010. *C. fragilis* subsp. *cavolinii* (Ten.) Damboldt, BG Bonn 25832 [Italy, Abruzzo], *Korotkova 81* (BONN), CAM099, FN397011. *C. glomerata* L., Germany, Hesse, Taunus, Helbigshainer Wiesen, *Borsch 3828* (B), CAM013, FN397012. *C. grossheimii* Kharadze, Georgia, *Schewardnadze s.n.* (BONN), CAM029, FN397013. *C. hierapetrae* Rech. f., BG Berlin 001- 03-96-44 [Krete, Sitias, Mt. Afendis Kavousi], *Gartenherbar 36642* [orig. leg. Turland] (B), CAM124, FN397014. *C. hohenackeri* Fisch. & al., Georgia, Tbilisi, Plato nutsubidze, *Schewardnadze 26* (B), CAM044, FN397015. *C. hypopolia* Trautv., Georgia, Mtiuletien, Tschkiri, *Otschiauri & Kimeridze s.n.* (TGM), CAM041, FN397016. *C. incurva* A. DC., BG Berlin 005-97-78-44 [Greece, Kifisia, Pilio], *Gartenherbar 10312* [leg. Broussalis s.n.] (B), CAM128, FN397017. *C. isophylla* Moretti, BG Berlin 35-10-74-80 [without locality data], *Gartenherbar 46378* (B), CAM127, FN397018. *C. jacobaea* C. Sm. ex Hook., BG Bonn 27130 [Cape Verde, Santiago, Serra Malagueta, road from Tarrafel to Assomada], CAM101, FN397019. *C. kolenatiana* C. A. Mey. ex Rupr., Georgia, *Schewardnadze s.n.* (BONN), CAM030, FN397020. *C. lactiflora* M. Bieb., Georgia, Kartli, road from Zagrevi to Bakuriani, *Korotkova 35* (B), CAM048, FN397021. *C. latifolia* L., BG Bonn 21712 [Germany, North-Rhine Westfalia, Eifel], *Lobin s.n.* (BONN), CAM059, FN397022. *C. makaschvilii* E. Busch, BG Bonn [Georgia], *Korotkova 98* (BONN), CAM032, FN397023. *C. medium* L., Switzerland, Lugano, *Borsch 3343* (B), CAM034, FN397024. *C. patula* L., BG Bonn Erhaltungskultur [Germany, North-Rhine Westfalia, Siegaue], *Borsch s.n.* (B, BONN), CAM073, FN397025; Slovakia, *Borsch & Košuth 3904* (B), CAM077, FN397026. *C. persicifolia* L., BG Bonn 16638 [Germany, North-Rhine Westfalia, Eifel, N of Gerolstein], *Finkel & Lobin s.n.* (BONN), CAM006, FN397027; Slovakia, *Borsch & Košuth 3912* (B), CAM081, FN397028. *C. petrophila* Rupr., Georgia, Kartli, Mt Gergeti, *Korotkova 21* (B), CAM050, FN397029. *C. portenschlagiana* Schult., BG Bonn 25634 [Croatia, Biokovo], *Korotkova 82* (BONN), CAM109, FN397030. *C. pyramidalis* L., BG Bonn 25636 [Croatia, Velebit], *Korotkova 83* (BONN), CAM108, FN397031. *C. rapunculoides* L., Germany, Baden-Württemberg, Hegau, Rosenegg, *Borsch s.n.* (B), CAM036, FN397032. *C. rhomboidalis* L., BG Oldenburg 03-116-06-10 [France, SW Alps, Lautaret], *Metzing s.n.,* CAM072, FN397033. *C. rotundifolia* L., Germany, North-Rhine Westfalia, Ahrtal, *Korotkova 2* (BONN), CAM008, FN397034. *C. rotundifolia* L. s.l., Italy, Lombardia, Monte Tremalzo, *Borsch 3854* (B), CAM086, FN396986. Italy, Lombardia, Mte Bazena, *Borsch 3856* (B), CAM084, FN396985. *C. rumeliana* (Hampe) Vatke, Greece, Mt Orvilos, *Raus & al. 21276*

(B), CAM133, FN397035. *C. sarmatica* Ker Gawl., Georgia, Mtiuletien, Sortscho, *Otschiauri & Kimeridze s.n.* (TGM), CAM042, FN397036. *C. sartorii* Boiss. & Heldr., BG Berlin 035-15-74-74 [Greece], *Gartenherbar 20042* (B), CAM126, FN397037. *C. saxifraga* M. Bieb., Georgia, Kartli, *Korotkova & al. 18* (B, TGM), CAM093, FN397038. *C. scheuchzeri* Vill., Germany, Bavaria, Karwendel Mts, *Borsch 3818* (B), CAM011, FN397039. *C. sibirica* subsp. *divergentiformis* (Jáv.) Domin, Slovakia, *Borsch & Košuth 3906* (B), CAM079, FN397042. *C. spatulata* Sm., Greece, Karditsa, *Willing 163.005-031* (B), CAM132, FN397040. *C. spicata* L., Italy, Lago di Garda, Valle Toscolano, *Borsch 3844* (B), CAM053, FN397044. *C. thessala* Maire, BG Berlin 266-52-80-14 [Greece, Larisa, Mt Ossa], *Gartenherbar 46375* (B), CAM129, FN397045. *C. thyrsoides* L., BG Bonn 26252 [orig. leg. Binder & al. s.n.; without locality data], *Korotkova 84* (BONN), CAM103, FN397046; BG Oldenburg 03-101-01-07 [ex BG Genf, without locality data], *Metzing s.n.* (B), CAM106, FN397047. *C. tomentosa* Lam., BG Bonn 17139 [without locality data] CAM023, FN397048. *C. trachelium* L., BG Bonn 24887 [Germany, North-Rhine Westfalia, Eifel], *Lobin & Möseler s.n.* (BONN), CAM007, FN397049; Italy, Lombardia, Mte Tremalzo, *Borsch 3855* (B), CAM83, FN397050. *C. trichocalycina* Ten., Greece, Karditsa, *Willing 142.206-218* (B), CAM130, FN397051. *C. tridentata* Schreb., BG Bonn 26253 [without locality data], *Korotkova 85* (BONN), CAM104, FN397052. *C. turczaninovii* Fed., BG Bonn [Mongolia, Arraykheer, NW Dugi, c. 60 km from Naiman Nuur in Urwurchangai- Aimak; orig. leg Tschentke], *Borsch 3965* (B), CAM074, FN397053. *C. versicolor* Andrews, BG Bonn 22978 [Greece, Acheron], *Korotkova 91* (BONN), CAM005, FN397054. *C. sp.,* Georgia, Kartli, road from Bakuriani to Zchrazcharo, *Korotkova & al. 38* (B), CAM094, FN397043. *Canarina canariensis* (L.) Vatke, BG Bonn 11906 [Canary Islands, Gran Canaria, Los Tilos de Moya], *Korotkova 8* (BONN), CAM018, FN397055. *C. eminii* Asch. & Schweinf., BG Bonn 8749 [Zaire, Parc National de Kahuzi-Biega, Tshivanga], *Fischer 1433q* (BONN), N454, FN397056. *Codonopsis lanceolata* (Siebold & Zucc.) Trautv., BG Bonn 19097 [without locality data], *Korotkova 92,* CAM064, FN397057. *Cyananthus lobatus* Wall. ex Benth., BG Bonn 25641 [Nepal, Langtan], no voucher, CAM065, FN397058. *Cyclocodon lancifolium* Kurz, BG Bonn 20135 [Laos, Luang Namtha], *Neumann s.n.* (B, BONN), N394, FN397059. *Edraianthus pumilio* (Schult.) A. DC., BG Bonn 25648 [Croatia, Biokovo], *Korotkova 97,* CAM063, FN397063. *E. tenuifolius* (Waldst. & Kit.) A. DC., BG Bonn 27244 [without locality data], *Korotkova 93* (BONN), CAM117, FN397064. *Jasione laevis* Lam., BG Bonn 19098 [Germany, Baden-Württemberg, S-Schwarzwald], *Korotkova 94* (BONN), CAM113, FN397069. *J. montana* L., France, Hérault, Seranne

Mts, *Borsch PII-138* (B), CAM040, FN397070. ***Legousia speculum-veneris*** (L.) Chaix, BG Bonn 7924 [without locality data], *Korotkova 86* (BONN), CAM114, FN397071. ***Michauxia campanuloides*** L'Hér., BG Bonn 25755 [without locality data], *Korotkova 87* (BONN), CAM118, FN397076. ***Musschia aurea*** Dum., BG Bonn 26210 [Portugal, Madeira, Punta do Garajau], Photo (BONN), CAM 112, FN397077. ***M. wollastoni*** Lowe, BG Bonn 22591 [without locality data], *Korotkova 67* (BONN), CAM022, FN397078. ***Nesocodon mauritianus*** (I. B. K. Richardson) Thulin, BG Bonn 11668 [Mauritius, Cascade], *Korotkova 7* (BONN), CAM24, FN397079. ***Petromarula pinnata*** (L.) A. DC., BG Bonn 24360 [Crete], *Korotkova 95* (BONN), CAM027, FN397081. ***Physoplexis comosa*** (L.) Schur, BG Bonn 24368 [Italy], Photo (BONN), CAM002, FN397082. ***Phyteuma betonicifolium*** Vill., Italy, Lombardia, Mte Bazena, *Borsch 3858* (B), CAM087, FN397084. ***P. hemisphaericum*** L., Switzerland, Gotthard, *Borsch 3869* (B), CAM058, FN397085. ***P. orbiculare*** L., France, Hérault, Seranne Mts, *Borsch 3697* (B), CAM009, FN397083. ***P. scheuchzeri*** All., Italy, Lombardia, Mte Tremalzo, *Borsch 3853* (B), CAM055, FN397086. ***Platycodon grandiflorum*** A. DC., BG Bonn 3387 [without locality data], *Korotkova 96* (BONN), CAM003, FN397087. ***Symphyandra hofmannii*** Pant., BG Bonn ?16577 [without locality data], CAM004, FN397092. ***Wahlenbergia saxicola*** A. DC., BG Bonn 4233 [without locality data], *Borsch s.n.* (BONN), CAM012, FN397093.

Campanulaceae-Cyphioideae: Cyphia peteriana E. Wimm., BG Bonn [without locality data], CAM 069, FN397060. ***C. subtubulata*** E. Wimm., BG Bonn [without locality data], CAM115, FN397061. ***C. tysonii*** Phillips, BG Bonn [without locality data], CAM116, FN397062.

Campanulaceae-Lobelioideae: Brighamia insignis A. Gray, BG Bonn 20142 [Hawaii], Photo (BONN), CAM25, FN396981. ***Grammatotheca bergiana*** C. Presl, BG Bonn 23854 [without locality data], *Korotkova 10* (BONN), CAM015, FN397066. ***Isotoma axillaris*** Lindl., BG Bonn 2104 [Australia], *Korotkova 6* (BONN), CAM020, FN397067. ***I. fluviatilis*** F. Muell. ex Benth., BG Bonn 587 [without locality data], *Korotkova 65* (BONN), CAM021, FN397068. ***Lobelia deckenii*** Hemsl., BG Bonn 17642 [Tansania, Mt. Meru], Photo (BONN), N439, FN397072. ***L. inflata*** L., USA, Virginia, Patrick County, Blue Ridge Mts, *Borsch & Neinhuis 3285* (B), CAM035, FN397073. ***L. nana*** Kunth, Argentina, Salta, San Antionio de los Cobres, *Borsch & al. 3744* (B, LPB), CAM028, FN397074. ***L. rhynchopetalum*** Hemsl., BG Bonn 00589 [without locality data], Photo (BONN), CAM014, FN397075. ***L. salicina*** Lam., Dominican Republic, *Borsch 3840* (B), CAM051, FN397065. ***Pratia nummularia*** A. Braun & Asch., BG Bonn 24780 [ex BG Regensburg, without locality data], *Korotkova 64* (BONN), CAM017, FN397088. ***Siphocampylus manettiiflorus*** Hook., BG Bonn 11102 [Cuba], *Korotkova 9* (BONN), CAM016, FN397089.

Appendix 2. List of indels found in the *petD* region of *Campanulaceae*.

The *petB-petD* spacer

1. "TATAG" SSR, in *Campanula elatines* and *C. portenschlagiana;* duplication from the gene end of *petB*.
2. A 9 nt gap in all members of the *Phyteuma* clade, probably deletion.
3. "TA"-SSR in all members of the lobelioids.
4. A 5 nt gap in *Platycodon grandiflorum,* probably deletion.
5. One-nucleotide gap in all other taxa except *Legousia speculum-veneris;* probably insertion of "G" in *Legousia,* creating a satellite-like "AG" repeat sequence.
6. One-nucleotide gap in all members of the core campanuloids and the *Wahlenbergia* clade, probably deletion.
7. "T" or "G" present in all members of the *Campanula rotundifolia* clade except *Adenophora triphylla* and *C. turczaninovii,* probably insertion.
8. One-nucleotide gap in *Lobelia inflata,* probably deletion.
9. One-nucleotide gap in all three species of *Cyphia,* probably deletion.
10. One-nucleotide gap in *Campanula turczaninovii* and *Adenophora,* probably deletion.
11. "CAAGTATGGATTATTAACAATAATAC" SSR in *Cyananthus lobatus*.
12. A 10 nt gap in *Pratia* and the two species of *Isotoma,* probably deletion.
13. One-nucleotide gap in *Grammatotheca,* probably deletion.
14. "TTATAT" SSR in the two species of *Jasione*.
15. "ATATATTATTTGATAT" in *Cyananthus,* probably insertion but origin not clear.
16. "TATTTTTTATATTC" SSR in *Campanula trichocalycina*.
17. A 4 nt gap in *Isotoma arillaris,* probably deletion.
18. "C" SSR in *Adenophora*.
19. One-nucleotide gap in *Lobelia nana,* probably deletion.

The *petD* intron

20. A 2 nt gap in *Campanula barbata* (CAM085 and CAM052), *C. hypopolia* and in *Edraianthus pumilio,* probably deletion.

21. A 5 nt gap in *Campanula carpatica*, probably deletion.

22. A 5 nt gap in both species of *Canarina*, probably deletion.

23. A one-nucleotide gap in *Campanula patula* (CAM073 and CAM077) and in *C. spatulata,* probably deletion.

24. "GTA" SSR in *Campanula hypopolia*.

25. A one-nucleotide gap in CAM051 *Lobelia salicina*.

25. A 2 nt gap in both species of *Jasione,* probably deletion.

26. A 3 nt gap in all species of *Cyphia,* probably deletion.

27. "A" SSR in *Edraianthus tenuifolius*.

28. "T" SSR in *Isotoma axillaris*.

29. "TGAT[A/C]" SSR in *Grammatotheca*.

30. "TGATA" SSR in all species of *Cyphia*.

31. A one-nucleotide gap in *Musschia aurea,* overlapping with 32, probably deletion.

32. A 3 nt gap in *Asyneuma canescens* and *Edraianthus pumilio,* overlapping with 31, probably deletion.

33. A 4 nt gap in *Adenophora triphylla* and *Campanula turczaninovii,* probably deletion.

34. "C" SSR in *Brighamnia, Campanula patula* (CAM073 and CAM077) and *C. spatulata*.

35. A 10 nt gap in *Cyphia tysonii,* probably deletion.

36. A 35 nt gap in *Brighamnia,* probably deletion.

37. A one-nucleotide gap in both species of *Edraianthus,* probably deletion.

38. "CCGAC" SSR in *Codonopsis* and *Cyananthus*.

39. "TTG" SSR in *Campanula baumgartenii*, *C. scheuchzeri*, *C. rhomboidalis* and *C. rotundifolia* (CAM008, CAM084, CAM086).

40. A 6 nt gap in *Edraianthus tenuifolius,* probably deletion.

41. A 11 nt gap in *Campanula trichocalycina,* probably deletion, overlapping with 43.

42. A 9 nt gap in both species of *Jasione,* probably deletion, overlapping with 43.

43. "GTCAK" in both species of *Musschia* and *Campanula lactiflora*, origin unknown.

44. A 4 nt gap in *Cyphia peteriana,* probably deletion.

45. "A" SSR in *Isotoma axillaris*.

46. "TAAGA" SSR in CAM051 *Lobelia salicina*.

47. A 7 to 9 nt gap in *Wahlenbergia,* overlapping with 48 to 49.

48. A 2 nt gap present in all lobelioids, cyphioids, *Canarina*, *Codonopsis*, *Cyananthus*, *Cyclocodon* and *Platycodon,* overlapping with 47 and 49.

49. A 4 nt gap in *Campanula trichocalycina,* overlapping with 48.

50. A 3 nt gap in *Campanula ciliata,* overlapping with 51 and 52.

51. A 4 nt gap in *Campanula carpatha* and *C. latifolia,* overlapping with 50.

52. A 5 nt gap in *Adenophora triphylla* and *Campanula turczaninovii,* overlapping with 50 and 51.

53. "GCKS" present in all taxa except core campanuloids, *Nesocodon* and *Wahlenbergia,* overlapping with 47.

54. Gap of approx. 51 nt in *Cyphia subtubulata,* probably deletion, overlapping with 55 to 65.

55. "TTACGTAAAA" SSR in *Cyananthus*.

56. "CCATTGCGTAAAACCATTACAGTAAAA" multiple SSR with partial repeats in *Codonopsis*.

57. A one-nucleotide gap in both species of *Jasione,* probably deletion, overlapping with 54.

58. "TC" SSR in *Cyananthus*.

59. An approx. 9 nt gap in *Brighamnia,* probably deletion, overlapping with 54 and 60.

60. An approx. 8 nt gap in *Lobelia deckenii* and *L. rhynchopetala,* probably deletion, overlapping with 54 and 59.

61. "TTCAGA" SSR in *Canarina eminii,* overlapping with 54.

62. A 9 nt gap in *Campanula caespitosa* and *C. cochleariifolia*, probably deletion, overlapping with 54.

63. A 2 nt gap in *Campanula latifolia* and *C.* sp. CAM094, probably deletion, overlapping with 54 and 62.

64. A 10 nt gap in *Lobelia nana,* overlapping with 54 and 65-66.

65. A 3 nt gap in *Lobelia inflata,* overlapping with 54 and 64.

66. "CGC" in *Lobelia inflata,* element of unknown origin.

67. A one-nucleotide gap in *Phyteuma betonicifolium,* probably deletion.

68. A 5 nt gap in both species of *Canarina,* probably deletion.

69. "TAGATA" SSR in *Asyneuma, Campanula trichocalycina*, *Legousia*, *Petromarula*, *Physoplexis* and all species of *Phyteuma*.

70. A one-nucleotide gap in *Campanula tomentosa,* probably deletion.

71. "A" in *Cyphia peteriana,* probably insertion, of unknown origin.

72. "G" SSR in both species of *Jasione*.

73. "A" SSR in *Grammatotheca*.

74. A 3 nt gap in *Campanula bravensis, C. edulis* and *C. jacobaea,* probably deletion.

Appendix 3. Positions of mutational hotspots in the individual sequences of *petD* that were excluded from phylogenetic analysis.

Taxon	Pos. hotspot 1	Pos. hotspot 2	Pos. hotspot 3
Outgroup taxa			
Atropa bella-donna	371-376	532-561	609-617
Lactuca sativa	335-343	496-531	594-594
Nicotiana tabacum	371-377	533-562	610-618
Pentaphragma acuminatum	381-388	549-583	625-628
Sphenoclea zeylanica	377-389	519-526	567-568
Stylidium majus	367-375	532-571	614-626
Campanulaceae-Campanuloideae			
Adenophora triphylla	334-344	504-523	569-571
Asyneuma campanuloides	338-348	513-532	584-586
Asyneuma canescens	335-345	510-529	581-583
Azorina vidalii	338-346	511-530	576-578
Campanula alliariifolia	338-346	511-530	576-578
Campanula alpina	338-346	511-530	576-578
Campanula annae	338-346	511-530	576-578
Campanula armazica	338-346	511-530	576-578
Campanula barbata	336-344	509-528	574-576
Campanula barbata	336-344	509-528	574-576
Campanula baumgartenii	338-348	516-535	581-583
Campanula betulifolia	338-346	511-530	576-578
Campanula bononiensis	338-346	511-530	576-578
Campanula bravensis	338-346	511-530	576-578
Campanula aff. *carnica*	338-348	516-535	581-583
Campanula carpatha	338-346	507-526	572-574
Campanula carpatica	333-343	508-527	573-575
Campanula cervicaria CAM037	338-346	511-530	576-578
Campanula cervicaria CAM078	338-346	511-530	576-578
Campanula cervicaria CAM082	338-346	511-530	576-578
Campanula cespitosa	337-347	512-531	568-570
Campanula choruhensis	338-346	511-530	576-578
Campanula ciliata	338-346	508-527	573-575
Campanula cochleariifolia	338-348	513-532	569-571
Campanula collina	338-346	511-530	576-578
Campanula darialica	338-346	511-530	576-578
Campanula divaricata	338-348	513-532	578-580
Campanula edulis	338-346	511-530	576-578
Campanula elatinoides	338-348	513-532	578-580
Campanula fedorovii	338-346	511-530	576-578
Campanula fenestrellata	338-348	513-532	578-580
Campanula foliosa	338-346	511-530	576-578
Campanula fragilis	338-348	513-532	578-580
Campanula glomerata	338-346	511-530	576-578
Campanula grossheimii	338-346	511-530	576-578
Campanula hierapetrae	338-346	511-530	576-578
Campanula hohenackeri	338-346	511-530	576-578
Campanula hypopolia	339-347	512-531	577-579
Campanula incurva	338-346	511-530	576-578
Campanula isophylla	338-348	513-532	578-580
Campanula jacobaea	338-346	511-530	576-578
Campanula kolenatiana	338-346	511-530	576-578
Campanula lactiflora	338-348	518-537	583-585
Campanula latifolia	338-346	507-526	570-572
Campanula makaschvili	338-346	511-530	576-578
Campanula medium	338-346	511-530	576-578
Campanula patula CAM073	338-348	513-532	578-580
Campanula patula CAM077	338-348	513-532	578-580
Campanula persicifolia CAM006	338-348	513-532	578-580
Campanula persicifolia CAM081	338-348	513-532	578-580
Campanula petrophila	338-346	511-530	576-578
Campanula portenschlagiana	338-348	513-532	578-580

Taxon	Pos. hotspot 1	Pos. hotspot 2	Pos. hotspot 3
Campanula pyramidalis	338-348	513-532	578-580
Campanula rapunculoides	338-346	511-530	576-578
Campanula rhomboidalis	338-348	516-535	581-583
Campanula rotundifolia CAM008	338-348	516-535	581-583
Campanula rotundifolia s.l. CAM084	338-348	516-535	581-583
Campanula rotundifolia s.l. CAM086	338-348	516-535	581-583
Campanula rumeliana	338-346	511-536	582-584
Campanula sarmatica	338-346	511-530	576-578
Campanula sartorii	338-346	511-530	576-578
Campanula saxifraga	338-346	511-530	576-578
Campanula scheuchzeri	338-348	516-535	581-583
Campanula sibirica subsp. *divergentiformis*	338-346	511-530	576-578
Campanula spatulata	338-348	513-532	578-580
Campanula sp. CAM094	338-346	511-530	574-576
Campanula spicata	338-346	511-530	576-578
Campanula thessala	338-346	511-530	576-578
Campanula thyrsoides CAM103	338-346	511-530	576-578
Campanula thyrsoides CAM106	338-346	511-530	576-578
Campanula tomentosa CAM023	338-346	511-530	576-578
Campanula trachelium CAM007	338-346	511-530	576-578
Campanula trachelium CAM083	338-346	511-530	576-578
Campanula trichocalycina	338-348	498-517	569-571
Campanula tridentata	338-346	511-530	576-578
Campanula turczaninovii	334-342	502-521	567-569
Campanula versicolor	338-348	513-532	578-580
Canarina canariensis	333-343	510-535	576-578
Canarina eminii	333-343	510-535	582-584
Codonopsis lanceolata	338-348	520-546	619-621
Cyananthus lobatus	338-348	520-545	602-604
Cyclocodon lancifolium	338-348	515-539	585-587
Edraianthus pumilio	333-341	505-524	570-572
Edraianthus tenuifolius	339-347	505-524	570-572
Jasione laevis	336-344	500-519	564-566
Jasione montana	336-344	500-519	564-566
Legousia speculum-veneris	338-346	511-530	582-584
Michauxia campanuloides	338-346	511-530	576-578
Musschia aurea	337-347	517-536	582-584
Musschia wollastonii	338-348	518-537	583-585
Nesocodon mauritianus	338-348	509-528	574-576
Petromarula pinnata	338-348	513-533	586-588
Physoplexis comosa	338-348	513-532	584-586
Phyteuma aff. *orbiculare*	338-350	515-534	586-588
Phyteuma betonicifolium	338-350	515-534	585-587
Phyteuma hemisphaericum	338-350	515-534	586-588
Phyteuma scheuchzeri	338-350	515-534	586-588
Platycodon grandiflorum	338-348	515-541	587-589
Symphyandra hofmannii	338-346	511-526	572-574
Wahlenbergia saxicola	338-344	504-523	570-572
Campanulaceae-Cyphioideae			
Cyphia peteriana	340-350	513-530	576-578
Cyphia subtubulata	340-351	518-529	533-535
Cyphia tysonii	340-350	507-524	570-572
Campanulaceae-Lobelioideae			
Brighamia insignis	339-347	479-503	540-542
Grammatotheca bergiana	343-351	518-540	586-588
Isotoma axillaris	339-347	515-539	585-587
Isotoma fluviatilis	338-346	513-535	581-584
Lobelia deckenii	338-346	513-537	575-577
Lobelia inflata	338-346	513-537	583-585
Lobelia nana	338-346	513-536	572-574
Lobelia rhynchopetalum	338-346	513-537	575-578
Lobelia salicina	337-345	517-541	587-588
Pratia nummularia	338-346	513-537	583-585
Siphocampylus manettiiflorus	338-346	513-537	583-586