TÜBİTAK

Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2014) 38: 203-212

doi:10.3906/bot-1308-46

Testing the utility of *matK* and ITS DNA regions for discrimination of *Allium* species

Meryem İPEK[1,*], Ahmet İPEK[1], Philipp W. SIMON[2]

[1]Department of Horticulture, Faculty of Agriculture, Uludağ University, Bursa, Turkey

[2]Department of Horticulture and US Department of Agriculture Agricultural Research Service, Vegetable Crops Research Unit, University of Wisconsin, Madison, USA

Received: 25.08.2013 • **Accepted:** 27.11.2013 • **Published Online:** 17.01.2014 • **Printed:** 14.02.2014

Abstract: Molecular phylogenetic analysis of the genus *Allium* L. has been mainly based on the nucleotide sequences of the ITS region. In 2009, *matK* and *rbcL* were accepted as a 2-locus DNA barcode to classify plant species by the Consortium for the Barcode of Life Plant Working Group. The *matK* region was chosen as a DNA barcode because of its effective species discrimination power, high quality sequence recovery, and easy experimental procedures. Integration of *matK* sequences into *Allium* phylogeny could improve phylogenetic reconstruction of this genus. This study was carried out to test the utility of nucleotide sequences of *matK* for discrimination of *Allium* species and to compare topologies of the phylogenetic trees based on *matK* and ITS analyses. Topologies of the phylogenetic trees based on ITS and *matK* analyses were very similar but a few accessions were placed into distant phylogenetic groups. Neither ITS nor *matK* analyses were able to discriminate some closely related *Allium* species alone. However, we do not suggest the use of a concatenated data approach to increase resolving power of ITS and *matK* because of the presence of the paralogous sequences and different types of cytoplasm in different accessions of a species. Therefore, we can suggest use of the *matK* region as an additional tool for phylogenetic analysis in *Allium* because characterization of the nucleotide sequences of *matK* region was easier to recover and more cost-effective than those of the ITS region.

Key words: *Allium*, barcode DNA, phylogenetic analysis, species differentiation

1. Introduction

About 750 species have been recognized within the genus *Allium* L. (Stearn, 1992) and at least 20 of these species are edible (Fritsch and Friesen, 2002). Economically imported and widely cultivated edible *Allium* species are onion (*A. cepa* L.), garlic (*A. sativum* L.), leek (*A. porrum* L.), chive (*A. schoenoprasum* L.), Chinese chive (*A. tuberosum* Rottl. ex Spreng), rakkyo (*A. chinense* G.Don) and Japanese bunching onion (*A. fistulosum* L.). Remaining edible *Allium* species are either cultivated locally or collected from wild for local consumption.

The phylogeny of the genus *Allium* has been revised using several molecular approaches. The initial molecular approach for phylogenetic analyses of the genus *Allium* was based on restriction enzyme analysis of chloroplast DNA (Havey, 1991; von Berg et al., 1996). In these studies, the phylogenetic relationships among *Allium* species were found to be generally in agreement with their previous traditional taxonomic classification. Later nucleotide sequences of the some conserved DNA regions were utilized for the phylogenetic reconstruction of the genus *Allium*. The nucleotide sequence of the nuclear ribosomal internal transcribed spacer region (ITS) was one of the most commonly used DNA regions for discrimination of the *Allium* species (Dubouzet and Shinoda, 1998, 1999; Mes et al., 1999; Friesen et al., 2000a; Fritsch and Friesen, 2002; Friesen et al., 2006; Gurushidze et al., 2007, 2008; Ipek et al., 2008). The ITS region is composed of the 5.8S ribosomal subunit and flanking internal transcribed spacers 1 and 2 (ITS-1 and ITS-2) of nuclear ribosomal DNA (nrDNA). The genus *Allium* was divided into 14 subgenera based on ITS sequences (Fritsch and Friesen, 2002). However, in a recent classification of the genus *Allium* with ITS data, 15 subgenera were recognized (Friesen et al., 2006). Besides the ITS region, the nucleotide sequences of chloroplast *rps16* (Li et al., 2010), *matK*, *trnH-psbA* (Son et al., 2010), *trnL-F*, *trnD-T* (Hirschegger et al., 2010), and *rpL32-trnL* (Wheeler et al., 2013) were also utilized for phylogenetic analysis of the genus *Allium*.

Recently, some selected DNA regions from chloroplast and nuclear genome have been proposed for barcoding plant species (reviewed in Hollingsworth et al., 2011). Nucleotide sequences of barcode DNA regions are intended to reveal polymorphisms mostly at interspecific

* Correspondence: msipek@uludag.edu.tr

or higher taxonomic levels but not at the intraspecific level (Kress and Erikson, 2007). With this, a barcode DNA region could discriminate species from each other and should clarify subgenera and sections within a genus. In this respect, nucleotide sequences of *matK*, *rpoC1 rpoB*, *trnH-PsbA rbcL*, *atpF-atpH*, and *psbK-psbI* and their combinations were recently tested for barcoding plant species by the Consortium for the Barcode of Life (CBOL) Plant Working Group (Hollingsworth et al., 2009) and *matK* and *rbcL* were accepted as a 2-locus DNA barcode. The *matK* DNA region has been tested for discrimination of many plant species as a single barcode region or in combination with other proposed plant barcode sequences in recent studies (Lahaye et al., 2008; Burgess et al., 2011; De Mattia et al., 2011; Du et al., 2011; Gu et al. 2011; Guo et al., 2011; Li et al., 2011; Pang et al., 2011; Seberg et al., 2012; Bandara et al., 2013). *matK* as a barcode DNA has advantages of effective species discrimination power, high quality sequence recovery, easy experimental procedures and sequence alignment, and lack of allelic polymorphisms or multiple paralogous copies compared to nuclear DNA regions (Lahaye et al., 2008; Hollingsworth et al., 2009; Burgess et al., 2011).

The integration of *matK* sequences into *Allium* phylogeny could improve phylogenetic reconstruction of this genus. However, applicability of the nucleotide sequence of *matK* for phylogenetic analysis in *Allium* was tested in only a single study using several species (Son et al., 2010). The purpose of the current study was to test the utility of nucleotide sequences of *matK* for discrimination of *Allium* species and to compare phylogenetic tree topologies of *matK* and ITS analyses.

2. Materials and methods

2.1. Plant materials

For both ITS and *matK* analyses, *Allium* species from 12 sections in 8 subgenera [based on the passport information in the Germplasm Resources Information Network (GRIN) of the US Department of Agriculture Agricultural Research Service (USDA-ARS)] were used for phylogenetic analyses. Nucleotide sequences of *matK* from 63 accessions and nucleotide sequences of ITS from 62 accessions in 32 *Allium* species were characterized for phylogenetic analysis (Table). Accessions of *A. cepa*, *A. tuncelianum*, *A. sativum* (cv. Kastamonu), and *A. porrum* (Porrum-1) were collected from Turkey. The rest of the accessions were obtained from the US Department of Agriculture, Western Region Plant Introduction Station, Pullman, WA, USA (Table).

Table. Accession number, subgenus, section, and origin of *Allium* species analyzed using the nucleotide sequences of ITS and *matK* DNA regions.

Accession	Species	Subgenus/Section	Origin
PI 405035	*A. vavilovii* Popov & Vved.	*Cepa/Cepa*	Former Soviet Union
PI 406677	*A. vavilovii* Popov & Vved.	*Cepa/Cepa*	Former Soviet Union
PI 576960	*A. vavilovii* Popov & Vved.	*Cepa/Cepa*	Former Soviet Union
PI 292163	*A. oschaninii* O.Fedtsch.	*Cepa/Cepa*	Former Soviet Union
W6 12754	*A. oschaninii* O.Fedtsch.	*Cepa/Cepa*	Kazakhstan
Cepa-1	*A. cepa* L.	*Cepa/Cepa*	Local market, Bursa, Turkey
Cepa-2	*A. cepa* L.	*Cepa/Cepa*	Local market, Bursa, Turkey
PI 592999	*A. roylei* Stearn	*Polyprason/Oreiprason*	Russian Federation
W6 12755	*A. pskemense* B.Fedtsch.	*Cepa/Cepa*	Uzbekistan
PI 576923	*A. pskemense* B.Fedtsch.	*Cepa/Cepa*	Netherlands
PI 576903	*A. galanthum* Kar. & Kir.	*Cepa/Cepa*	Former Soviet Union
W6 17644	*A. galanthum* Kar. & Kir.	*Cepa/Cepa*	California, USA
PI 280549	*A. altaicum* Pall.	*Cepa/Cepa*	Former Soviet Union
PI 576874	*A. altaicum* Pall.	*Cepa/Cepa*	Nei Monggol, China
PI 219754	*A. fistulosum* L.	*Cepa/Cepa*	Tokyo, Japan
PI 369186	*A. fistulosum* L.	*Cepa/Cepa*	Former Soviet Union
PI 576875	*A. altyncolicum* N.Friesen	*Cepa/Schoenoprasum*	Altai, Russian Federation
PI 371880	*A. schoenoprasum* L.	*Cepa/Schoenoprasum*	Former Soviet Union
PI 401719	*A. schoenoprasum* L.	*Cepa/Schoenoprasum*	Denmark
PI 369185	*A. obliquum* L.	*Polyprason/Oreiprason*	Former Soviet Union

Table. (Continued).

Accession	Species	Subgenus/Section	Origin
PI 576906	*A. hymenorrhizum* Ledeb.	*Polyprason/Falcatifolia*	Siberia, Russian Federation
PI 485591	*A. hymenorrhizum* Ledeb.	*Polyprason/Falcatifolia*	Sinkiang, China
PI 576882	*A. amphibolum* Ledeb.	*Reticulatobulbosa/ Reticulatobulbosa*	Altai, Russian Federation
PI 576913	*A. lineare* L.	*Reticulatobulbosa/ Reticulatobulbosa*	Siberia, Russian Federation
W6 21059	*A. lineare* L.	*Reticulatobulbosa/ Reticulatobulbosa*	Altai, Russian Federation
W6 18960	*A. mongolicum* Regel	*Rhizirideum/Caespitosoprason*	Mongolia
W6 18966	*A. polyrhizum* Turcz. ex Regel	*Rhizirideum/Caespitosoprason*	Mongolia
PI 369526	*A. senescens* L.	*Rhizirideum/Rhizirideum*	Former Soviet Union
PI 261804	*A. senescens* subsp. *montanum* (Fr.) Holub	*Rhizirideum/Rhizirideum*	Austria
PI 369525	*A. angulosum* L.	*Rhizirideum/Rhizirideum*	Former Soviet Union
PI 576917	*A. nutans* L.	*Rhizirideum/Rhizirideum*	Altai, Russian Federation
W6 9997	*A. nutans* L.	*Rhizirideum/Rhizirideum*	Montana, USA
PI 576912	*A. libani* Boiss.	*Melanocrommyum/Acanthoprason*	New York, USA
PI 576937	*A. rotundum* subsp. *rotundum*	*Allium/Allium*	New York, USA
Tuncelianum-1	*A. tuncelianum* (Kollman) Özhatay et al.	*Allium/Allium*	Tunceli, Turkey
Tuncelianum-2	*A. tuncelianum* (Kollman) Özhatay et al.	*Allium/Allium*	Tunceli, Turkey
Tuncelianum-3	*A. tuncelianum* (Kollman) Özhatay et al.	*Allium/Allium*	Tunceli, Turkey
Tuncelianum-4	*A. tuncelianum* (Kollman) Özhatay et al.	*Allium/Allium*	Tunceli, Turkey
PI 483425	*A. porrum* L.	*Allium/Allium*	Former Soviet Union
Porrum-1	*A. porrum* L.	*Allium/Allium*	Local market, Bursa, Turkey
PI 207457	*A. ampeloprasum* L.	*Allium/Allium*	Kabul, Afghanistan
W6 14000	*A. ampeloprasum* L.	*Allium/Allium*	Denmark
PI 167255	*A. ampeloprasum* L.	*Allium/Allium*	Turkey
PI 576936	*A. scorodoprasum* L.	*Allium/Allium*	Bulgaria
DDR7116	*A. sativum* L.	*Allium/Allium*	Germany
PI 493118	*A. sativum* L.	*Allium/Allium*	Skierniewice, Poland
cv. Kastamonu	*A. sativum* L.	*Allium/Allium*	Local market, Kastamonu, Turkey
PI 515774	*A. sativum* L.	*Allium/Allium*	Togo
PI 540370	*A. sativum* L.	*Allium/Allium*	California, USA
PI 383817	*A. sativum* L.	*Allium/Allium*	Former Serbia and Montenegro
W6 1903 (U094)	*A. longicuspis* L.	*Allium/Allium*	Uzbekistan
DDRGRU2	*A. sativum* L.	*Allium/Allium*	Germany
DDR6811*	*A. sativum* L.	*Allium/Allium*	Germany
PI 615416 (U079)	*A. sativum* L.	*Allium/Allium*	Uzbekistan
W6 1961	*A. sativum* L.	*Allium/Allium*	Cuenca, Spain
PI 280557	*A. tuberosum* Rottler ex Spreng.	*Butomissia/Butomissia*	Former Soviet Union
PI 576957	*A. tuberosum* Rottler ex Spreng.	*Butomissia/Butomissia*	Bishkek, Kyrgyzstan
PI 399173	*A. tuberosum* Rottler ex Spreng.	*Butomissia/Butomissia*	New Territories, Hong Kong
PI 264799	*A. ramosum* L.	*Butomissia/Butomissia*	Switzerland
PI 371878	*A. ramosum* L.	*Butomissia/Butomissia*	Former Soviet Union
PI 576926	*A. ramosum* L.	*Butomissia/Butomissia*	Russian Federation
W6 20304	*A. cernuum* Roth	*Amerallium/Lophioprason*	Norway
PI 372503	*A. cernuum* Roth	*Amerallium/Lophioprason*	Ontario, Canada

*: Analyzed by using only *matK* region.

2.2. Plant propagation and DNA sampling

Seeds from each accession, or cloves of clonally propagated *Allium sativum* accessions, were planted into 2-L pots. Leaf samples were collected from 1–3 individual plants from each accession at 3–5 leaf stages and they were freeze-dried for 3 days. DNA samples were extracted from 100 mg of lyophilized and powdered leaf samples using a modified CTAB method and the concentrations of DNA samples were adjusted to 30 ng/µL (Ipek et al., 2003).

2.3. ITS analysis

ITS regions (ITS-1, 5.8S rDNA subunit, ITS-2) were amplified using a primer combination of ITSA and ITSB (Blattner, 1999). Each 20-µL polymerase chain reaction (PCR) contained 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.4 µM of each primer, 1 U of Taq DNA polymerase (PanVera, Madison, WI, USA) with supplied reaction buffer at 1X concentration, and 45 ng of template DNA. Thermal cycle conditions for the ITS region were performed according to protocol described by Ipek et al. (2008). PCR products of the ITS region were size fractionated by electrophoresing through 1.5% (w/v) agarose gel in 1X TAE buffer (40 mM Tris–acetate, pH 8.0, and 1 mM EDTA). Gels were stained with ethidium bromide (0.5 mg mL^{-1}) (Sigma, St Louis, MO, USA) and visualized digitally. PCR-amplified DNA fragments of the ITS region were excised and eluted from agarose gels and cloned into the pGEM–T Easy Vector (Promega, Madison, WI, USA) by using procedures described by Ipek et al. (2005, 2006). Three bacterial colonies from each purified and cloned PCR product were subjected to PCR amplification using T7 and SP6 universal primers of the cloning vector using the same reaction and thermal cycling condition described above. PCR products diluted 10 times with sterile ddH2O were used for sequencing reactions.

2.4. *matK* analysis

Polymerase chain reactions for the *matK* region contained the same components as described above. Thermal cycle conditions for the *matK* region were carried out in accordance with the method described by Cuénoud et al. (2002) using the same primer pair (390F 5'-CGATCTATTCATTCAATATTTC-3' and 1326R 5'-TCTAGCACACGAAAGTCGAAGT-3'). Briefly, reactions were heated to 94 °C for 1 min for initial denaturation and exposed to 26 cycles of 94 °C for 1 min, 48 °C for 30 s, and 72 °C for 1 min, with a final extension of 7 min at 72 °C. PCR-amplified DNA fragments of the *matK* region were directly used in sequencing reactions after 10× dilution with sterile ddH2O.

2.5. Sequencing

Taq DNA polymerase (Applied Biosystems, Foster City, CA, USA) cycle-sequencing reactions were performed according to the conditions recommended by Applied Biosystems with reagents using a fluorescent-dye terminator, but using half-volume reactions. Products of cycle sequencing reactions were run on a PE-Biosystem 377 XL (Applied Biosystems) automated DNA sequencing instrument at the Biotechnology Center of the University of Wisconsin, Madison, WI, USA. Both ITS and *matK* regions were sequenced in both directions. For all PCR reactions, the GeneAmp PCR System 2700 (Applied Biosystems) was used.

2.6. Data analysis

Sequences of ITS and *matK* regions were manually edited with CHROMAS v.2.31 (Technelysium Pty. Ltd., South Brisbane, Australia) and aligned by using CodonCode Aligner 3.7.1 software (CodonCode Corporation, Centerville, MA, USA). No manual adjustment in the alignment of the nucleotide sequences for the ITS region was made to avoid introducing subjective bias.

Aligned ITS and *matK* sequences were evaluated with bootstrap analysis (2000 replicates) (Felsenstein, 1985) and Kimura (1980) distances were calculated to construct a neighbor-joining (NJ) dendrogram using the TREECON v. 1.3b program (van de Peer and de Wachter, 1994) with the program defaults. The NJ trees were visualized using TREECON.

Aligned ITS and *matK* sequences were also used for the maximum parsimony (MP) method. MP trees were obtained using the tree-bisection-regrafting (TBR) algorithm (Nei and Kumar, 2000) with search level 1, in which the initial trees were obtained by the random addition of sequences (10 replicates). MP trees were tested with bootstrap analysis (1000 replicates). Branch lengths of MP trees were calculated using the average pathway method (Nei and Kumar, 2000) and are in the units of the number of changes over the whole sequence. All positions containing gaps in ITS sequences were eliminated. MP analyses were conducted using the program MEGA5 (Tamura et al., 2011).

The nucleotide sequences of ITS region (gi:11595756) (Friesen et al., 2000b) and *matK* (gi:345132216) (Shackelford and Fay, 2012) of *Nothoscordum bivalve* L. were obtained from the GenBank databases at the National Center for Biotechnology Information (NCBI) and were included in the analyses for rooting the NJ and MP trees.

3. Results

The NJ tree developed using Kimura (1980) distances based on nucleotide sequences of the ITS region had a similar topology as the MP tree, shown in Figure 1. Phylogenetic tree based on NJ analysis of Kimura (1980) distances of the nucleotide sequences of the *matK* region was topologically similar to MP tree in Figure 2. However, trees generated by NJ and MP analyses differed in their branch-length values (data not presented).

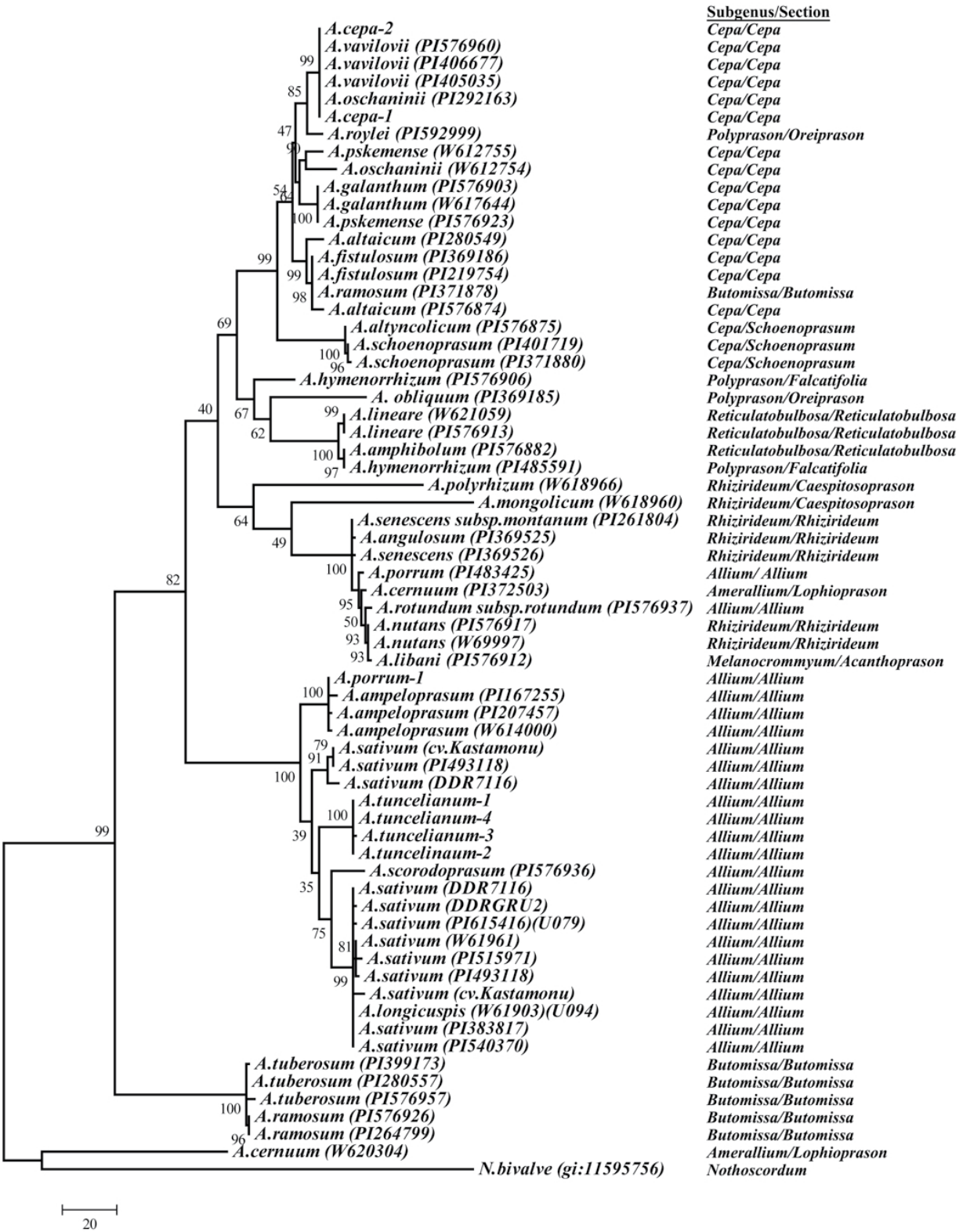

Figure 1. Phylogenetic analysis of 32 *Allium* species based on the nucleotide sequences of ITS region. Most parsimonious tree was developed using the TBR algorithm (Nei and Kumar, 2000). Numbers next to the branches are the bootstrap test values. The scale is the branch lengths calculated using the average pathway method (Nei and Kumar, 2000).

Phylogenetic analyses based on the nucleotide sequences of ITS and *matK* generally discriminated subgenera and sections in the genus *Allium*, although some accessions of a species were placed outside of the section (Figures 1 and 2). For example, 2 accessions belong to section *Allium*, PI 483425 (*A. porrum*) and PI 576937 (*A. rotundum* subsp. *rotundum*), were placed outside of the section *Allium* clade.

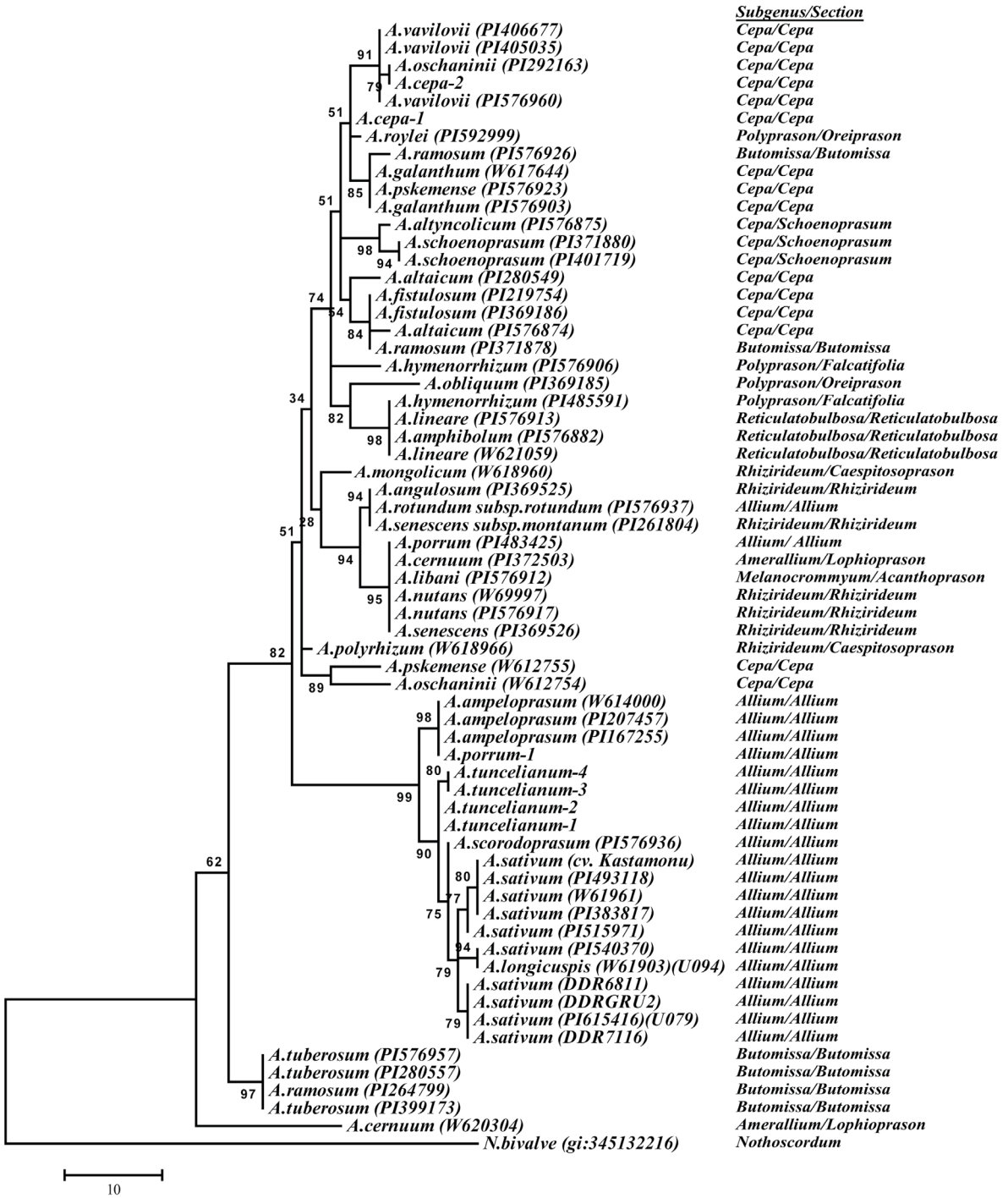

Figure 2. Phylogenetic analysis of 32 *Allium* species based on the nucleotide sequences of *matK* region. Most parsimonious tree was developed using the TBR algorithm (Nei and Kumar, 2000). Numbers next to the branches are the bootstrap test values. The scale is the branch lengths calculated using the average pathway method (Nei and Kumar, 2000).

Although the ITS region separated most of the species in the genus *Allium* from each other, some closely related species could not be discriminated (Figure 1). For example, *A. vavilovii* Popov & Vved., *A. cepa*, and 1 accession of *A. oschaninii* O.Fedtsch. were located in the same clade without any differences. Phylogenetic analysis of the ITS region also demonstrated that, for several species, accessions were located in distant clades, suggesting that there might be mixtures. For example, although 2 accessions of *A. ramosum* L. were located in a clade with the accessions of *A. tuberosum* Rottler ex Spreng., another accession (PI 371878) of *A. ramosum* was located in a distant clade with *A. altacium* Pall. and *A. fistulosum* L. (Figure 1). To check for possible mixture during the analysis, DNA samples from 2 more independent plants of PI 371878 (*A. ramosum*) were analyzed and the results were same. In another case, 1 accession of *A. hymenorrhizum* Ledeb. (PI 485591) was sister to accessions of *A. amphibolum* Ledeb. and *A. lineare* L. but another accession of *A. hymenorrhizum* (PI 576906) was more distantly related (Figure 1). All accessions of *A. sativum* were located in the same clade but the presence of paralogous sequences of the ITS region were detected in 3 *A. sativum* accessions (DDR7116, PI 493118, and cv. Kastamonu) (Figure 1).

Although most of the species were also discriminated by *matK*, some closely related species could not be distinguished (Figure 2). For example, accessions of *A. tuberosum* and *A. ramosum* were placed in the same clade with no clear differences. *matK* analysis also confirmed mixtures. For instance, different accessions of *A. cernuum* Roth were placed in distant clades. A DNA sample from 1 more independent plant of PI 372503 (*A. cernuum*) was also analyzed to check for mixture during the experimental procedures, but the result was same.

Topologies of the phylogenetic trees based on *matK* and ITS regions were similar, but there were some *Allium* species that were placed in different clades with ITS or *matK* analysis. For example, ITS analysis placed *A. ramosum* accessions PI 576926 and PI 264799 into the clade of section *Butomissa* with the accession of *A. tuberosum* (Figure 1). However, *mat*K analysis placed PI 576926 into the clade of section *Cepa* (Figure 2). Similarly, an accession (W6 12755) of *A. pskemense* and an accession (W6 12754) of *A. oschaninii* were placed into the clade of section *Cepa* by ITS analysis (Figure 1), but these accessions were clustered in a phylogenetically distant group by *matK* analysis (Figure 2). These differences between ITS and *matK* analyses were probably due to the interspecific hybridization between the *Allium* species.

4. Discussion

Previously, the ITS region was suggested for phylogenetic analysis in plant species (Baldwin et al., 1995; Dubouzet and Shinoda, 1999; Chen et al., 2010; Dirmenci et al., 2010; Dündar et al., 2013). Recently, nucleotide sequences of some DNA regions in the plastid genome (*matK*, *rpoC1 rpoB*, *trnH-PsbA rbcL*, *atpF-atpH*, *psbK-psbI*) and their combinations were tested for barcoding plant species. Among these DNA regions, *matK* and *rbcL* were accepted as a 2-locus DNA barcode by the CBOL Plant Working Group (Hollingsworth et al., 2009). Polymorphisms in barcode nucleotide sequences at interspecific or higher taxonomic levels are desired, but not at the intraspecific level. In this way, an unknown plant accession can be assigned to a species by comparing the nucleotide sequences of the barcode DNA region (Hebert et al., 2003; Stoeckle, 2003; Kress and Erikson, 2008; Pang et al., 2011). In addition, the DNA region for barcoding should be highly recoverable, should have a high percentage of species resolution, and should be cost-effective (Burgess et al., 2011). In this study, utility of the nucleotide sequences of *matK* was tested as a DNA barcode for discrimination of the *Allium* species and was compared with the previously suggested ITS region. According to our results, ITS and *matK* regions were both easy to amplify with PCR. Species resolution of both DNA regions was similar because the topologies of trees of ITS and *matK* were similar. Both DNA regions generally separated sections in *Allium*. However, several species could not be discriminated from each other by either ITS or *matK* regions. On the other hand, the ITS region requires cloning before sequencing because of the allelic polymorphisms (presence of insertion/deletion, single nucleotide polymorphism), pseudogenes, and paralogous copies of the ITS region in a plant species (Buckler et al., 1997; Denduangboripant and Cronk, 2000; Kita and Ito, 2000; Bailey et al., 2003; Li et al., 2004; Razafimandimbison et al., 2004; Ruggiero and Procaccini, 2004; King and Roalson, 2008). In our study, the nucleotide sequences of ITS ranged from 720 bp in *A. pskemense* (W6 12754) to 748 bp in *A. cernuum* (W6 20304). Allelic polymorphisms and paralogous copies of the ITS region complicate sequence characterization without cloning. However, the requirement for cloning increases costs and labor significantly for sequence characterization of ITS region. On the other hand, there was no allelic polymorphism or insertion/deletion in the *matK* region within a plastid genome of *Allium* species in this study and the size of the *matK* region was 857 bp in all *Allium* species. Therefore, we were able to efficiently amplify and sequence-characterize the *matK* region without cloning.

Genetic relationships among 9 *Allium* species were recently analyzed using the polymorphisms within the nucleotide sequences of *trnH-psbA*, ITS, and *matK* (Son et al., 2010). The authors reported the presence of more than 1 haplotype in both organellar and nuclear genomes in a single plant of the *Allium* species, which makes phylogenetic analysis and barcoding impossible using the

nucleotide sequences of these DNA regions. In contrast, we observed only 1 haplotype in each plant of *Allium* species for *matK* in the plastid genome. Although it is not common, there were 2 paralogous copies of the ITS region in a single plant of 3 accessions of *A. sativum* (DDR7116, PI 493118, and cv. Kastamonu) (Figure 1).

Phylogenetic relationships among *Allium* species based on both ITS and *matK* analyses were in agreement with the previous studies using the ITS region and restriction enzyme analysis of the chloroplast genome (Havey, 1992; Gurushidze et al., 2007; Li et al., 2010). Previously, *A. vavilovii* and *A. cepa* were grouped together without any polymorphism by Havey (1992) and closely clustered by Gurushidze et al. (2007). Similarly, *A. cepa* and *A. vavilovii* were closely clustered with both ITS and *matK* analyses in this study. An accession of *A. oschaninii* (PI 292163) was also closely clustered with *A. cepa* and *A. vavilovii*, but another accession of *A. oschaninii* (W6 12754) was closely grouped with *A. pskemense*. Accessions of *A. oschaninii* and *A. pskemense* were more closely related in the study of Gurushidze et al. (2007). Although *A. roylei* Stearn was classified under *Allium* subgenus *Polyprason* section *Oreiprason* according to the passport information in the GRIN of the USDA-ARS, this species was closely grouped with other *Allium* species in subgenus *Cepa* section *Cepa* with both ITS and *matK* analyses. The same result was also obtained in previous studies (Havey, 1992; Gurushidze et al., 2007; Li et al., 2010).

Two accessions of *A. cepa* were placed into the same clade by ITS analysis (Figure 1). On the other hand, with *matK* analysis, 1 accession of *A. cepa* was closely related to *A. oschaninii* but another accession of this species was located in a sister cluster with *A. roylei* (Figure 2). In *A. cepa*, cytoplasmic genetic male sterility was utilized for hybrid seed production. In this male sterility system, expression of male sterility requires sterile cytoplasm (S or T). Havey (1993) suggested that sterile cytoplasm might be an alien cytoplasm that could be the result of interspecific hybridization. This accession was obtained from a local market in Bursa, Turkey, and could be a hybrid onion cultivar with sterile cytoplasm. Indeed, PCR analyses with 2 different PCR-based markers developed by Sato (1998) and Engelke et al. (2003) demonstrated that this *A. cepa* accession had "S" cytoplasm (data not presented). In addition, these markers were also amplified within the genome of *A. roylei* and the sizes of the PCR-amplified DNA fragment were similar to that of "S" cytoplasm in *A. cepa*. PCR-amplified DNA fragments of these markers in "S" cytoplasm of *A. cepa* and *A. roylei* were sequence-characterized. Comparison of the nucleotide sequences of the PCR-amplified DNA fragments of these markers demonstrated the presence of small insertion/deletion and single nucleotide polymorphisms. Our observation of an *A. cepa* accession that was phylogenetically more closely related to *A. roylei* confirms the possibility of the foreign origin of "S" cytoplasm of *A. cepa*, and the progenitor species of "S" cytoplasm could be a species phylogenetically closely related to *A. roylei*.

Our results suggested that *Allium* species in subgenera *Polyprason* and *Reticulatobulbosa* were closely related. Similarly, in a previous study, it was demonstrated that species in these subgenera had close phylogenetic relationships (Li et al., 2010). All accessions of *Allium* species in subgenus *Rhizirideum* section *Rhizirideum* were clustered in the same clade (Figures 1 and 2). However, 2 accessions from subgenus *Allium* section *Allium*, 1 accession from subgenus *Melanocrommyum* section *Acanthoprason*, and 1 accession from subgenus *Amerallium* section *Lophioprason* were also grouped in the clade of subgenus *Rhizirideum* section *Rhizirideum* with both ITS and *matK* analyses. This result indicates the presence of mixtures in these accessions. Accessions belonging to subgenus *Allium* section *Allium* were clustered in a clade and supported by a 100% bootstrap value. Three accessions of *A. tuberosum* and 2 accessions of *A. ramosum* in subgenus *Butomissa* section *Butomissa* were clustered in a distant clade. Because of the great morphological similarity between *A. tuberosum* and *A. ramosum*, Hanelt (2001) merged both into *A. ramosum*, but Blattner and Friesen (2006) proposed to keep both taxa as separate species. Although our results based on the *mat*K and ITS regions also suggested that these species were phylogenetically closely related, analysis of ITS region discriminated these 2 species, but there was no polymorphism between the *matK* regions of *A. ramosum* and *A. tuberosum* (Figures 1 and 2). *A. cernuum* in subgenus *Amerallium* section *Lopioprasum* was the most distant species from the rest of the *Allium* species analyzed in this study.

In conclusion, based on ITS and *matK* analyses, phylogenetic relationships among *Allium* species analyzed in this study were in agreement with previous studies (Havey, 1992; Gurushidze et al., 2007; Li et al., 2010). However, the presence of mixtures in some accessions of *Allium* species analyzed was observed. Although the topology of phylogenetic trees based on ITS and *matK* analyses were similar in this study and ITS has served well for phylogenetic analysis for the genus *Allium*, we suggest use of *matK* as an additional tool for phylogenetic analysis in *Allium* because characterization of the nucleotide sequences of the *matK* region was easier to recover and more cost-effective than those of ITS region.

Acknowledgment

The authors thank Dr David M Spooner (Professor of Horticulture at the University of Wisconsin, Madison, WI, USA) for his critical review of the manuscript.

References

Bailey CD, Carr TG, Harris SA, Hughes CE (2003). Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol Phylogenet Evol 29: 435–455.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995). The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann Mo Bot Gard 82: 247–277.

Bandara NL, Papini A, Mosti S, Brown T, Smith LM J (2013). A phylogenetic analysis of genus *Onobrychis* and its relationships within the tribe *Hedysareae* (Fabaceae). Turk J Bot 37: 981–992.

Blattner FR (1999). Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. Biotechniques 27: 1180–1186.

Blattner FR, Friesen N (2006). Relationships between Chinese chive (*Allium tuberosum*) and its putative progenitor *A. ramosum* as assessed by random amplified polymorphic DNA (RAPD). In: Zeder MA, Bradley DG, Emshwiller E, Smith BD, editors. Documenting Domestication: New Genetic and Archaeological Paradigms. Berkeley, CA, USA: University of California Press, pp. 134–142.

Buckler ES, Ippolito A, Holtsford TP (1997). The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145: 821–832.

Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC, Newmaster SG, Percy DM, Hajibabaei M, Barrett SCH (2011). Discriminating plant species in a local temperate flora using the *rbcL+matK* DNA barcode. Method Ecol Evol 2: 333–340.

Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, Zhu YJ, Ma XY, Gao T, Pang XH et al. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5: e8613.

Cuénoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase MW (2002). Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL, atpB, and matK* DNA sequences. Am J Bot 89: 132–144.

De Mattia F, Bruni I, Galimberti A, Cattaneo F, Casiraghi M, Labra M (2011). A comparative study of different DNA barcoding markers for the identification of some members of Lamiacaea. Food Res Int 44: 693–702.

Denduangboripant J, Cronk QCB (2000). High intraindividual variation in internal transcribed spacer sequences in *Aeschynanthus* (Gesneriaceae) implications for phylogenetics. P Roy Soc Lond B Bio 267: 1407–1415.

Dirmenci T, Dündar E, Deniz G, Arabacı T, Martin E, Jamzad Z (2010). Morphological, karyological and phylogenetic evaluation of *Cyclotrichium*: a piece in the tribe Mentheae puzzle. Turk J Bot 34: 159–170.

Du ZY, Qimike A, Yang CF, Chen JM, Wang QF (2011). Testing four barcoding markers for species identification of Potamogetonaceae. J Syst Evol 49: 246–251.

Dubouzet JG, Shinoda K (1998). Phylogeny of *Allium* L. subg. *Melanocrommyum* (Webb et Berth.) Rouy based on DNA sequence analysis of the internal transcribed spacer region of rDNA. Theor Appl Genet 97: 541–549.

Dubouzet JG, Shinoda K (1999). Relationships among old and new world *Alliums* according to ITS DNA sequence analysis. Theor Appl Genet 98: 422–433.

Dündar E, Akçiçek E, Dirmenci T, Akgün Ş (2013). Phylogenetic analysis of the genus *Stachys* sect. *Eriostomum* (Lamiaceae) in Turkey based on nuclear ribosomal ITS sequences. Turk J Bot 37: 14–23.

Engelke T, Terefe D, Tatlioglu T (2003). A PCR-based marker system monitoring CMS-(S), CMS-(T) and (N)-cytoplasm in the onion (*Allium cepa* L.). Theor Appl Genet 107: 162–167.

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

Friesen N, Fritsch RM, Blattner FR (2006). Phylogeny and new intrageneric classification of *Allium* (Alliaceae) based on nuclear ribosomal DNA ITS sequences. Aliso 22: 372–395.

Friesen N, Fritsch RM, Pollner S, Blattner FR (2000a). Molecular and morphological evidence for an origin of the aberrant genus *Milula* within Himalayan species of *Allium* (Alliaceae). Mol Phylogenet Evol 17: 209–218.

Friesen N, Fritsch RM, Pollner S, Blattner FR (2000b). *Nothoscordum bivalve* 5.8S rRNA gene, internal transcribed spacer 1 (ITS1) and internal transcribed spacer 2 (ITS2). Bethesda, MD, USA: GenBank Databases at NCBI. Website http://www.ncbi.nlm.nih.gov/nuccore/gi:11595756 [accessed 14 January 2013].

Fritsch RM, Friesen N (2002). Evolution, domestication, and taxonomy. In: Rabinowitch, HD, Currah L, editors. *Allium* Crop Science: Recent Advances. New York, NY, USA: CABI Publishing, pp. 5–30.

Gu J, Su JX, Lin RZ, Li RQ, Xiao PG (2011). Testing four proposed barcoding markers for the identification of species within *Ligustrum* L. (Oleaceae). J Syst Evol 49: 213–224.

Guo X, Simmons MP, But PPH, Shaw PC, Wang RJ (2011). Application of DNA barcodes in *Hedyotis* L. (Spermacoceae, Rubiaceae). J Syst Evol 49: 203–212.

Gurushidze M, Fritsch RM, Blattner FR (2008). Phylogenetic analysis of *Allium* subgen. *Melanocrommyum* infers cryptic species and demands a new sectional classification. Mol Phylogenet Evol 49: 997–1007.

Gurushidze M, Mashayekhi S, Blattner FR, Friesen N, Fritsch RM (2007). Phylogenetic relationships of wild and cultivated species of *Allium* section *Cepa* inferred by nuclear rDNA ITS sequence analysis. Plant Syst Evol 269: 259–269.

Hanelt, P. 2001. Alliaceae. In: Hanelt P, Institute of Plant Genetics and Crop Plant Research, editors. Mansfeld's Encyclopedia of Agricultural and Horticultural Crops. Berlin, Germany: Springer-Verlag, pp. 2250–2269.

Havey MJ (1991) Phylogenetic relationships among cultivated *Allium* species from restriction enzyme analysis of the chloroplast genome. Theor Appl Genet 81: 752–757.

Havey MJ (1992) Restriction enzyme analysis of the chloroplast and nuclear 45S ribosomal DNA of *Allium* sections *Cepa* and *Phyllodolon* (Alliaceae). Plant Syst Evol 183: 17–31.

Havey MJ (1993). A putative donor of S-cytoplasm and its distribution among open-pollinated populations of onion. Theor Appl Genet 86: 128-134.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003). Biological identifications through DNA barcodes. P R Soc Lond B 270: 313–321.

Hirschegger P, Jaksě J, Trontelj P, Bohanec B (2010). Origins of *Allium ampeloprasum* horticultural groups and a molecular phylogeny of the section *Allium* (*Allium*: Alliaceae). Mol Phylogenet Evol 54: 488–497.

Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL, Fazekas AJ et al. (2009). A DNA barcode for land plants. P Natl Acad Sci USA 106: 12794–12797.

Hollingsworth PM, Graham SW, Little DP (2011). Choosing and using a plant DNA barcode. PLoS One 6: e19254.

Ipek M, Ipek A, Almquist SG, Simon PW (2005). Demonstration of linkage and development of the first low-density genetic map of garlic based on AFLP markers. Theor Appl Genet 110: 228–236.

Ipek M, Ipek A, Simon PW (2003). Comparison of AFLPs, RAPD markers, and isozymes for diversity assessment of garlic and detection of putative duplicates in germplasm collections. J Am Soc Hortic Sci 128: 246–252.

Ipek M, Ipek A, Simon PW (2006). Sequence homology of polymorphic AFLP markers in garlic (*Allium sativum* L.). Genome 49: 1246–1255.

Ipek M, Ipek A, Simon PW (2008). Genetic characterization of *Allium tuncelianum*: An endemic edible *Allium* species with garlic odor. Sci Hortic 115: 409–415

Kimura M (1980). A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111–120.

King MG, Roalson EH (2008). Exploring evolutionary dynamics of nrDNA in *Carex* subgenus *Vignea* (Cyperaceae). Syst Bot 33: 514–524.

Kita Y, Ito M (2000). Nuclear ribosomal ITS sequences and phylogeny in East Asian *Aconitum* subgenus *Aconitum* (Ranunculaceae), with special reference to extensive polymorphism in individual plants. Plant Syst Evol 225: 1–13.

Kress WJ, Erickson DL (2007). A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS One 2: e508.

Kress WJ, Erickson DL (2008). DNA barcodes: genes, genomics, and bioinformatics. P Natl Acad Sci USA 105: 2761–2762.

Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V (2008). DNA barcoding the floras of biodiversity hotspots. P Natl Acad Sci USA 105: 2923–2928.

Li FW, Kuo LY, Rothfels CJ, Ebihara A, Chiou WL, Windham MD, Pryer KM (2011). *rbcL* and *matK* earn two thumbs up as the core DNA barcode for Ferns. PLoS One 6: e26597.

Li J, Ledger J, Ward T, Del Tredici P (2004). Phylogenetics of Calycanthaceae based on molecular and morphological data, with a special reference to divergent paralogues of the nrDNA ITS region. Harv Pap Bot 9: 69–82.

Li QQ, Zhou SD, He XJ, Yu Y, Zhang YC, Wei XQ (2010). Phylogeny and biogeography *of Allium* (Amaryllidaceae: Allieae) based on nuclear ribosomal internal transcribed spacer and chloroplast *rps16* sequences, focusing on the inclusion of species endemic to China. Ann Bot-London 106: 709–733.

Mes THM, Fritsch RM, Pollner S, Bachmann K (1999). Evolution of the chloroplast genome and polymorphic ITS regions in *Allium* subgenus *Melanocrommyum*. Genome 42: 237–247.

Nei M, Kumar S (2000). Molecular Evolution and Phylogenetics. New York, NY, USA: Oxford University Press.

Pang X, Song J, Zhu Y, Xu H, Huang L, Chen S (2011). Applying plant DNA barcodes for Rosaceae species identification. Cladistics 27: 165–170.

Razafimandimbison SG, Kellogg EA, Bremer B (2004). Recent origin and phylogenetic utility of divergent ITS putative pseudogenes: a case study from Naucleeae (Rubiaceae). Syst Biol 53: 177–192.

Ruggiero MV, Procaccini G (2004). The rDNA ITS region in the Lessepsian marine angiosperm *Halophila stipulacea* (Forssk.) Aschers. (Hydrocharitaceae): Intragenomic variability and putative pseudogenic sequences. J Mol Evol 58: 115–121.

Sato Y (1998). PCR amplification of CMS-specific mitochondrial nucleotide sequences to identify cytoplasmic genotypes of onion (*Allium cepa* L.). Theor Appl Genet 96: 367–370.

Seberg O, Petersen G, Davis JI, Pires JC, Stevenson DW, Chase MW, Fay MF, Devey DS, Jørgensen T, Sytsma KJ et al. (2012). Phylogeny of the Asparagales based on three plastid and two mitochondrial genes. Am J Bot 99: 875–89.

Shackelford GE, Fay MF (2012). *Nothoscordum bivalve* voucher Chase 247, NCU maturase K (*matK*) gene, partial cds; chloroplast. Bethesda, MD, USA: GenBank Databases at NCBI. Website http://www.ncbi.nlm.nih.gov/nuccore/gi:345132216 [accessed 14 January 2013].

Son JH, Park KC, Kim TW, Park YJ, Kang JH, Kim NS (2010). Sequence diversification of 45S rRNA ITS, *trnH-psbA* spacer, and *matK* genic regions in several *Allium* species. Genes Genom 32: 165–172.

Stearn WT (1992). How many species of *Allium* are known? Kew Magazine 9: 180–182.

Stoeckle M (2003). Taxonomy, DNA and the bar code of life. Bioscience 53: 2–3.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.

van de Peer Y, de Wachter R (1994). TREECON for windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Biosci 10: 569–570.

von Berg GL, Samoylov A, Klaas M, Hanelt P 1996. Chloroplast DNA restriction analysis and infrageneric grouping of *Allium* (Alliaceae). Plant Syst Evol 200: 253–261. Wheeler EJ, Mashayekhi S, McNeal DW, Columbus JT, Pires JC (2013). Molecular systematics of *Allium* subgenus *Amerallium* (Amaryllidaceae) in North America. Am J Bot 100: 701–711.