AcademicPres

Available online: www.notulaebiologicae.ro

Print ISSN 2067-3205; Electronic 2067-3264

Not Sci Biol, 2016, 8(4):451-455. DOI: 10.15835/nsb.8.4.9884

Notulae Scientia Biologicae

Original Article

Sequence Analysis of the Internal Transcribed Spacer (ITS) Region of the Nuclear Ribosomal DNA (nrDNA) and Chloroplast *trn*L-F Region (cpDNA) of Some *Lactuca* L. (Asteraceae) Species in Turkey

Emre SEVİNDİK*, Veysel UZUN, Fatih COŞKUN

*Adnan Menderes University, Faculty of Agriculture, Department of Agricultural Biotechnology, South Campus Cakmar-Aydin, Turkey; ph.d-emre@hotmail.com (*corresponding author)*

Abstract

In the current study, sequence analysis of some Turkish *Lactuca* L. species using nrITS DNA and *trn*L-F cpDNA sequences were performed to elucidate phylogenetic relationships among the taxa under study. *Hieracium umbellatum* was used as an outgroup. Different plant materials of *Lactuca* were collected from different parts of Turkey during excursions of summer 2013. Plant materials were either kept in silica gel or kept fresh for immediate DNA isolation. Both phenol chloroform-isoamyl alcohol method and commercial kits were used to extract genomic DNA for PCR reactions. ITS4 and ITS5A primers were utilized for ITS region, while *trn*Le and *trn*Lf primers were used to amplify the *trn*L-F region. Obtained DNA sequences were edited both manually and by using BioEdit 7.0.4.1. Sequencing data were aligned via ClustalW program and analyzed using PAUP 4.01b10 software. nrITS sequences varied from 639 nucleotides to 735 nucleotides. Average nucleotide composition for nrITS was 22.1% (T), 27.9% (C), 23.2% (A) and 26.8% (G). It was also found that divergence values differed between 0.0000 and 0.10290. The *trn*L-F sequences varied from 296 nucleotides to 385 nucleotides. Average nucleotide composition of *trn*L-F sequences was 34.1% (T), 18.4% (C), 31.6% (A) and 16.0% (G). It was also found that divergence values differed between 0.0000 and 0.09674. Neighbour Joining (NJ) trees were constructed in order to identify the relationships among *Lactuca* species. Phylogenetic trees based on ITS region were found to be more useful than phylogenetic trees based on *trn*L-F region. After analysis of the results obtained, the data suggest that *Lactuca* contains 2 clades, with clade 1 having 2 subclades. These results support the prior phylogenetic studies on *Lactuca* and hence provide an up to date review of Turkish *Lactuca* species.

Keywords: *Lactuca*, ITS, *trn*L-F, sequence, Turkey

Introduction

The generic name *Lactuca* L. is derived from a Latin word "lac" means "milk", a common character to all the members of the tribe Cichorieae. The genus *Lactuca* L. (s.l) was established by Linnaeus (Linnaeus, 1753; Linnaeus, 1754). This genus belong to Asteraceae family which is represented by approximately 100 species and notably temperate and warm regions in Europe, Asia, North America, Australia and Africa widely distrubuted (Dziechciarková *et al*., 2004). The number of this species is 8 in Turkish flora (Davis, 1975). This number is given as 33 taxa according to the latest arrangements (Güner *et al.*, 2012). Species of this genus have important species for both economics and medicine. For example, *L. sativa* is used in salads since it contains high levels of vitamins and minerals. *L. serriola* seeds are used for treating insomnia and coughing. At the same time, they are used in manufacturing varnish, scented soaps and dying materials. Lactucarium, a medication obtained from *L. sativa* and *L. serriola* is used for treating bronchitis and asthma (Bano and Qaiser, 2011). *L. tatarica* (L.) C.A. is a significant forage plant, but when excessively consumed, becomes toxic (Kirpicznikov, 1964).

Internal transcribed spacer (ITS) regions are molecular markers used in molecular systematic classification of many plants (Baldwin *et al*., 1995; Ogundipe and Chase, 2008). ITS contains noncoding regions that are characterized by sometimes higher nucleotide substitution rates and have proven useful for elucidating phylogenetic relationships at the species and genus levels (Taberlat *et al*., 1991; Baldwin *et al*., 1993; Baldwin *et al*., 1995; Molvray *et al*., 1999; Ranjbar *et al*., 2014).

Chloroplast DNA (cpDNA) sequence variation is widely used in investigating the interspecific relation between

Received: 17 Aug 2016. Received in revised form: 17 Nov 2016. Accepted: 14 Dec 2016. Published online: 16 Dec 2016.

angiosperm and other plants. However, low evolutionary rate of these molecules is a limitation at interspecific level (Taberlat *et al.*, 1991; Türktaş *et al*., 2012).

Intergenic space is situated between *trn*L (UAA) 3'exon and *trn*F (GAA) gene (Taberlat *et al.*, 1991; Gielly and Taberlat, 1994; Gielly and Taberlat,, 1996; Liu *et al*., 2005). while *trn*L gen contains two protected exons. These are divided by a group of introns. This region has a high level of migration. For this reason, they are used in interspecies phylogenetic analyses (Yang and Pak, 2006).

In the hereby study, a molecular phylogenetic analysis of some Turkish *Lactuca* L. species using nrITS DNA and *trn*L-F cpDNA sequences were performed to elucidate the phylogenetic relationships using *H. umbellatum* as an outgroup.

Materials and Methods

Plant samples and DNA extractions

Lactuca species used in the study were collected from certain regions in Turkey in July and August 2013. Locations of these species are given in Appendix 1. Other taxa are *L. undulata, L. viminea, L. tatarica, L. orientalis, L. tuberosa* and *L. sativa*. ITS and *trn*L-F sequences of these taxa were acquired from NCBI genbank, and their genbank numbers are as follows. *L. undulata* (KF485647.1, KF486159.1), *L. orientalis* (KF485659.1, KF486171.1), *L. viminea* (AJ633333.1, KF486172.1), *L. sativa* (L13957.1, AY504775.1), *L. tuberosa* (only *trnL-F* KF486163.1) and *L. tatarica* (KF485661.1, KF486174.1). Total genomic DNAs were extracted from either silica gel dried leaves or from fresh materials following the methods from Dellaporta *et al.* (1983). The DNA pellet was dissolved in TE (tris HCL- EDTA) buffer. The purified total DNA was quantified by electrophoresis and its quality was determined by spectrophotometry. DNA samples were stored -20 °C.

PCR amplifications and sequencing

Double-stranded DNA of the complete ITS and *trn*L-F regions were amplified in each genomic DNA. Amplification of the whole ITS region (ITS1 + 5.8S + ITS2) was performed with primers ITS5A (Standford *et al*., 2000) and ITS4 (White *et al.*, 1990) (Table 1). The amplification process was performed in 25 µL of PCR reaction volume. Each PCR reaction contained 2.5 µL of Taq buffer, 1.5 µL of magnesium chloride ($MgCl_2$), 0.4 µL of dNTP, 2.5 µL for ITS4 and 2.5 µL for ITS5A primers, 0.3 µL of Taq DNA polymerase (Fermentas), 2,5 µL of total genomic DNA and 10.8 µL of ddH_2O. PCR amplification was performed with an initial denaturation step of 94 °C for 5 min, followed by 35 cycles of strand denaturation at 94 °C for 45 s, annealing at 50 °C for 45 s, and primer extension at 72 °C for 2 min, and a final elongation at 72 °C for 10 min. *trn*L-F moleculer marker analyzed in this study belongs to the chloroplast genome (cpDNA). Polymerase Chain Reaction (PCR) amplifications of *trn*L-F cpDNA were performed using the primers designed by Taberlet *et al.* (1991) (Table 1). The amplification process was performed in 25 µL of PCR reaction volume. Each PCR reaction contained 2.5 µL of total genomic DNA, 10.8 µL of ddH_2O, 2.5 µL of Taq buffer, 1.5 µL of Magnesium chloride ($MgCl_2$), 0.4 µL of dNTP, 2.5 µL for *trn*Le and 2.5 µL for *trn*Ff primers and 0.3 µL of Taq DNA polymerase. PCR amplification was performed with an initial denaturation step of 94 °C for 5 min, followed by 35 cycles of strand denaturation at 94 °C for 30 sec, annealing at 51 °C for 45 s, and primer extension at 72 °C for 90 s, and a final elongation at 72 °C for 8 min. Purified PCR products were sequenced with the amplifications primers by a commercial biotechnology company. For each sample, forward and reverse sequencing reactions were performed and the sequences checked via GenBank (NCBI) via blast search. Later obtained DNA sequences were edited both manually and by using the Bioedit 7.0.4.1 (Hall, 1999).

Aligment and phylogenetic analysis

nrITS and *trn*L-F sequences were aligned using ClustalW aligment software (Thompson *et al*., 1994). Ends of the alignment were trimmed to make all the sequences of equal length, which was a total of 744 nucleotide (nt) positions in the final dataset for ITS region and 440 for *trn*L-F. Aligned DNA sequences were turned to the NEXUS format for the phylogenetic analysis in the PAUP 4.01b10. Sequences distance values and Neighbour Joining trees were using PAUP 4.01b10 (Swofford, 2001). To evaluate the degree of support for given clades the bootstrap analysis (Felsenstein, 1985) was applied. Beside *H. umbellatum* (for ITS sequence HQ131822.1, as well as for *trn*L-F sequence KF196061.1) was used outgroup and retrieved from NCBI GenBank.

Results and Discussion

In this study, nrITS sequences ranged from 639 nt to 735 nt among 12 specimens under study. Average nucleotide composition was 22.1% (T), 27.9% (C), 23.2% (A) and 26.8% (G). The total length of the aligned ITS sequence matrix was of 745 nucleotides. There were a total of 105 variable characters of which 81 were parsimony informative and 559 characters were constant. Genetic distance method based on ITS sequence set was performed with PAUP 4.01b10 (Table 2). The lowest sequence divergence with ingroup taxa was noted between *L. kemaliya* and *L. serriola* with a value of 0.0000. The highest sequence divergence with ingroup taxa between *L. intricata* and *L. tatarica* was 0.10290 (Table 1).

ITS sequence based on NJ tree comprised of two main clade. Clade 1, consisted of eight taxa, while clade 2 consisted of three taxa. Clade 1 is divided into 2 sub-groups (A and B). Group A consisted of *L. viminea*, *L. orientalis* and *L. tatarica*. A quite high bootstrap value of 100% supported that *L. viminea* and *L. orientalis* are sister taxa (Fig. 1). Group B is a monophyletic group consisting of *L. aculeata*, *L. serriola*, *L. sativa, L. kemaliya* and *L. saligna*. Bootstrap analysis result supported this monophyletic group with a value of 91%. Within the branch, L. *aculeate*, *L. serriola, L. kemaliya* and *L. sativa* are sisters with a bootstrap value of 98% (Fig. 1). Clade 2 consisted of *L. undulata*, *L. incrinata* and *L. tuberosa.* Within the branch, *L. undulata* and *L.incrinata* are sister taxa and this relation is supported by a 100% bootstrap value. *L. tuberosa* was found to be closely related to this sister group with a bootstrap value of 95%.

Table 1. ITS and *trn*L-F primers used for the study

Primer name	5' to 3' Primer sequence	Based on (the source publication)
Forward ITS5A	CCTTATCATTTAGAGGAAGGAG	Stanford *et al*., (2000)
Reverse ITS4	TCCTCCGCTTATTGATATGC	White *et al*., (1990)
Forward trnLe	GGTTCAAGTCCCTCTATCCC	Taberlet *et al*., (1991)
Reverse trnFf	ATTTGAACTGGTGACACGAG	Taberlet *et al*., (1991)

Table 2. Pairwise sequence distances among some *Lactuca* species for ITS nrDNA data using PAUP distance matrix (except outgroup)

Taxa	1	2	3	4	5	6	7	8	9	10
L.orientalis	-									
L.viminea	0.00626	-								
L.tatarica	0.06566	0.06567	-							
Laculeata	0.05323	0.05323	0.07154	-						
L.serriola	0.05160	0.05160	0.07189	0.00288	-					
L.kemaliya	0.05138	0.05138	0.07129	0.00279	0.0000	-				
L.sativa	0.05974	0.06136	0.08321	0.01314	0.01027	0.01304	-			
L.saligna	0.06527	0.06528	0.07143	0.03236	0.03288	0.03213	0,04496			
L.undulata	0.09267	0.09110	0.09263	0.09083	0.09083	0.09044	0.10073	0.8740	-	
L.intricata	0.09546	0.09385	0.10290	0.08720	0.08672	0.08447	0.09794	0.08749	0.02968	-
L.tuberosa	0.09614	0.09936	0.10204	0.09154	0.09178	0.09120	0.09623	0.08980	0.07102	0.07563

Table 3. Pairwise sequence distances among some *Lactuca* species for *trn*L-F cpDNA data using PAUP distance matrix (except outgroup)

Taxa	1	2	3	4	5	6	7	8	9	10
L.tatarica	-									
L.tuberosa	0.01690	-								
L.undulata	0.02693	0.02696	-							
L.aculeata	0.01352	0.01733	0.02973	-						
L.kemaliya	0.01666	0.02044	0.03288	0.00538	-					
L.sativa	0.01374	0.01689	0.02348	0.00000	0.00311	-				
L.serriola	0.01374	0.01689	0.02348	0.00000	0.00311	0.00000	-			
L.intricata	0.02750	0.03074	0.04722	0.08866	0.09674	0.01419	0.01419	-		
L.saligna	0.02791	0.03073	0.04449	0.01389	0.02307	0.01433	0.01433	0.00000	-	
L.orientalis	0.03115	0.03109	0.04409	0.01731	0.02044	0.01722	0.01722	0.02223	0.02229	-
L.viminea	0.02730	0.03052	0.04013	0.01389	0.01705	0.01355	0.01355	0.03208	0.02889	0.00336

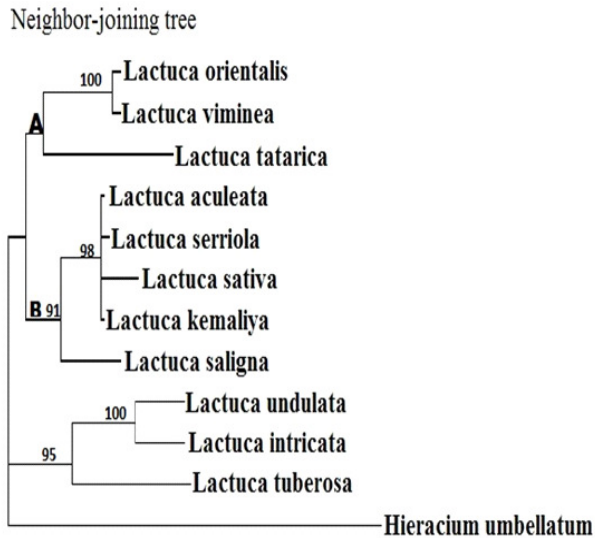

Fig. 1. The Neighbour Joining tree generated using ITS DNA sequences of genus *Lactuca* specimens and the related sequences retrieved from NCBI GenBank

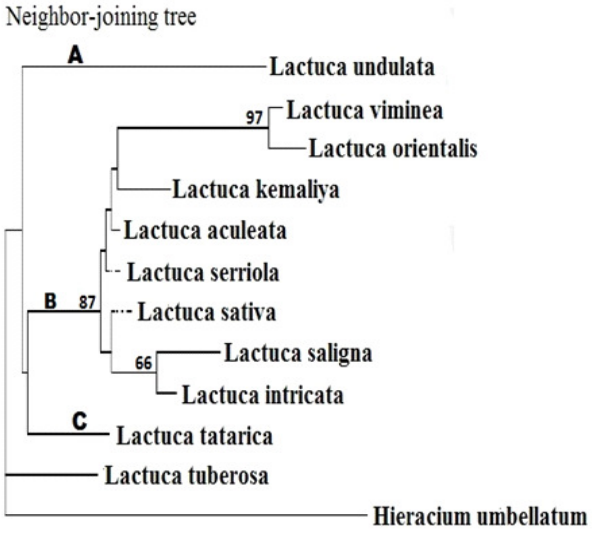

Fig. 2. The Neighbour Joining tree generated using *trn*L-F DNA sequences of genus *Lactuca* specimens and the related sequences retrieved from NCBI GenBank

For *trn*L-F sequences obtained a base lenght that ranged from 296 to 385 among the 12 specimens. Average nucleotide composition was of 34.1% (T), 18.4% (C), 31.6% (A) and 16% (G). The total length of the aligned *trn*L-F sequence matrix were 440 nucleotides. There were a total of 28 variable characters of which 7 were parsimony informative and 405 characters were constant. Genetic distance method based on *trn*L-F set was performed with PAUP 4.01b10 (Table 3). The lowest sequence divergence ingroup taxa was noted between *L. saligna* and *L. intricata*, between *L. sativa* and *L. aculeata*, between *L. serriola* and *L. aculeata*, between *L. serriola* and *L. sativa* respectivley 0,0000. The highest sequence divergence ingroup taxa between *L. intricata* and *L. kemaliya* was 0.09674. *trn*L-F sequence based on NJ tree comprised of two main clade (Fig. 2). Clade 1 consisted of ten taxa, while clade 2 consisted of one taxa.

NJ tree created based on *trn*L-F sequences was found to be slightly different than the NJ tree created based on ITS sequences. Clade 1 was divided into 2 sub-groups (A and B) within itself. Group A consists of only *L. undulata*. Group B is made up of *L. viminea, L. orientalis, L. kemaliya, L. aculeata, L. serriola, L. sativa, L. saligna* and *L. intricata,* and is supported by an 87% bootstrap value. Within the group, *L. viminea* and *L. orientalis* are sister taxa and were supported by bootstrap value of 97%. Similarly, *L. saligna* and *L. intricata* are sister taxa and were supported by a 66% bootstrap value. Group C is only made up of *L. tatarica* (Fig. 2). Clade 2, consist of only *L. tuberosa* (Fig. 2).

There are many phylogenetics and genetic variety studies conducted on *Lactuca* species (Landry *et al*., 1987; Kesseli *et al*., 1991; Kesseli *et al*., 1994; Vermeulen *et al.*, 1994; Koopman *et al.*, 2001). Koopman *et al*. (1998) carried out phylogenetic analysis of among *Lactuca* species based on ITS1 region. The results of their study were similar to ours. In the current study, ITS and *trn*L-F analysis indicated that *L. saligna, L. serriola*, *L. aculeata* and *L. sativa* were in the same clade. The results of a study by Koopman *et al.* (1998) were also similar to the hereby obtained data. In the present study, in ITS sequence, *L. tuberosa* was within clade 2 and was distinguished from *L. sativa, L. saligna, L. serriola L. viminea* and *L. tatarica.* A similar result was obtained in the study by Koopman *et al.,* (1998). The only difference noted between the two experiments was sequenced at ITS1 region. Whereas in the current study were sequenced all of ITS1, 5.8S and ITS2 regions. Therefore, it was obtained a longer base pair.

Conclusions

Recently, phylogenetic analyses based on DNA have been used for the purpose of revealing taxonomic data. These analyses are used in rebuilding phylogenetics of many taxonomically complex species, types or groups that contain a great number of taxons. Because of the fact that phylogeny is widely used, many methods were developed for reconstructing these. In the current study, sequence analysis of several Turkish *Lactuca* species, using ITS and *trn*L-F sequences, were performed to elucidate phylogenetic relationships. Thus, the study can be illustrated as a supportive molecular and phylogenetic research carried out on some *Lactuca* species of Anatolia.

Acknowledgements

The authors are very greateul to Prof. Dr. Kamil Coskunçelebı and Assist. Prof. Dr. M. Yavuz Paksoy for providing the plant materials for DNA extraction. This research was supported by Balıkesir University with grant number 2013/130.

References

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. Annals of the Missouri Botanical Garden 82:247-277.

Baldwin BG (1993). Molecular phylogenetics of *Calycadenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution reexamined. American Journal of Botany 80:222-238.

Bano R, Qaiser M (2011). Taxonomic revision of the genus *Lactuca* L. (Cichorieae- Asteraceae) from Pakistan and Kashmir. Pakistan Journal of Botany 43(5):2259-2268.

Davis PH (1975). Flora of Turkey and the East Aegean Island. Vol 5. Edinburgh Univ. Press.

Dellaporta SL, Wood J, Hicks JB (1983). A plant DNA minipreparation: Version II. Plant Molecular Biology Reporter 1:19-21.

Dziechciarková M, Lebeda A, Doležalová I, Astley D (2004). Characterization of *Lactuca* spp. germplasm by proteinand molecular markers – a review. Plant Soil and Environent 50(2):47-58.

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

Gielly L, Taberlet P (1994). The use of chloroplast DNA to resolce plant phylogenies: non-coding versus *rbcL* sequences. Molecular Biology and Evolution 11:769-777.

Gielly L, Taberlet P (1996). A phylogeny of the European gentians inferred from chloroplast *trnL* (UAA) intron sequences. Botanical Journal of the Linnean Society 120:57-75.

Güner A, Akyıldırım B, Alkayış MF, Çıngay B, Kanoğlu SS, Özkan AM, Öztekin M, Tuğ GN (2012). Türkçe bitki adları. Şu eserde: Güner A, Aslan S, Ekim T, Vural M, Babaç MT (Eds). Türkiye Bitkileri Listesi (Damarlı Bitkiler). Nezahat Gökyiğit Botanik Bahçesi Yayınları Flora Dizisi 1 İstanbul.

Hall TA (1999). Bioedit: a user-friendly biological sequence alignment editor and analyses program for Windows 95/98/NT. Nucleic Acids Symposium 41:95-98.

Kesseli R, Ochoa O, Michelmore R (1991). Variation at RFLP loci in *Lactuca* spp. and origin of cultivated lettuce (*L. sativa*). Genome 34:430-436.

Kesseli RV, Paran I, Michelmore RW (1994). Analysis of a detailed genetic linkage map of *Lactuca sativa* (Lettuce) constructed from RFLP and RAPD markers. Genetics 136:1435-1446.

Kirpicznikov ME (1964). In: Flora of the USSRi. Bobrov EG (Ed). Tzevlev. Compositae 29(7-9):255-355. Leningrad.

Koopman MJM, Guetta E, Van De Wiel MCC, Vosman B, Van Den Berg GR (1998). Phylogenetic relationships among *Lactuca* (Asteraceae) species and related genera based on ITS-1 DNA sequences. American Journal of Botany 85(11):1517-1530.

Koopman MJM, Zevenbergen JM, Van Den Berg GR (2001). Species relationships in *Lactuca* S.L. (Lactuceae, Asteraceae) ınferred from AFLP fingerprints. American Journal of Botany 88(10):1881-1887.

Landry BS, Kesseli RV, Farrara B, Michelmore RW (1987). A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331-337.

Linnaeus C (1754). Genera Plantarum, 5th ed. NewYork pp 815-816.

Linnaeus C (1753). Species Plantarum, 1st ed. The Ray Society, London pp 795-796.

Liu Q, Ge S, Tang H, Zhang X, Zhu G, Lu B-R (2005). Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ricosomal internal transcribed spacer and chloroplast *trn*L-F sequences. New Phytologist 170:411-420.

Molvray M, Kores JP, Chase WM (1999). Phylogenetic relationships within *Korthalsella* (Viscaceae) based on nuclear ITS and plastid *trn*L-F sequence data. American Journal of Botany 86(2):249-260.

Ogundipe OT, Chase M (2008). Phylogenetic analyses of Amaranthaceae based on *mat*K DNA sequence data with emphasis on West African species. Turkish Journal of Botany 33:153-161.

Ranjbar M, Hajmoradi F, Waycott M, Van Dijk K (2014). A phylogeny of the tribe Caraganeae (Fabaceae) based on DNA sequence data from ITS. Feddes Repertorium 125:78-84.

Stanford AM, Harden R, Parks CR (2000). Phylogeny and biogeography of *Juglans* (Juglandaceae) based on *mat*K *and* ITS sequence data. American Journal of Botany 87:872-882.

Swofford DL (2001). PAUP. Phylogenetic Analysis Using Parsimony (and other methods), Version 4.0b10. Sinaeur Associates, Sunderland.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17:1105-1109.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Symposium Series 22:4673-4680.

Türktaş M, Aslay M, Kaya E, Ertuğrul F (2012). Molecular characterization of phylogenetic relationships in *Fritillaria* species inferred from chloroplast *trn*L-*trn*F sequences. Turkish Journal of Biology 36:552-560.

Vermeulen A, Desprez B, Lancelin D, Bannerot H (1994). Relationship among *Cichorium* species and related genera as determined by analysis of mitochondrial RFLPs. Theoretical and Applied Genetics 88:159-166.

White TJ, Bruns TD, Lee SB, Taylor JW (1990). PCR protocols: a guide to methods and applications. San Diego (CA): Academic Press 315-322.

Yang YJ, Pak HJ (2006). Phylogeny of Korean *Rubus* (Rosaceae) Based on ITS(nrDNA) and *trn*L/F Intergenic Region (cpDNA). Journal of Plant Biology 49(1):44-54.