DNA Barcoding of Two Narrow Endemic Plants; Astragalus Argaeus and Astragalus Stenosemioides from Mount Erciyes, Turkey

handan Şapcı Selamoğlu (✉ handansapci@kayseri.edu.tr)

Kayseri University   https://orcid.org/0000-0001-8150-0450

Research Article

Abstract

The genetic resources and biological diversity of countries are very important. Biodiversity and genetic resources should be protected, especially as endemic species. In this concept, DNA barcoding studies are an effective way to identify an unknown taxon and protected the biodiversity of a country. *Astragalus argaeus* and *A. stenosemioides* are narrow endemic species from Mt. Erciyes, Turkey. To determine its phylogenetic relationships and DNA barcoding, sequence data from the chloroplast DNA (*matK* region) were analyzed by parsimony, maximum likelihood, and Bayesian methods. In this study, *A. argaeus, A. stenosemioides* samples, and 23 sequences from GenBank, including a closely related species were performed. The phylogenetic analysis showed that the *matK* gene region could clearly distinguish *A. argaeus* and *A. stenosemioides* from its closely related species. DNA barcoding surveys can contribute to taxonomic and biodiversity research, various molecular ecology, and population genetics studies. Also, it is possible to determine the species by separating the *matK* DNA gene region, which is one of the molecular characters, and *A. argaeus* and *A. stenosemioides* have been successfully barcoded for the first time; therefore, it has been shown that this gene region can be used for barcoding.

Introduction

Many molecular markers are used for plant and animal genetic studies. Recently, the most commonly used methods to determine the genetic variation of DNA-based methods such as DNA barcoding, phylogenetic studies (Carew *et al.* 2005). It is very important for taxonomy and biodiversity studies by using DNA barcode methods to fast and correctly make taxa identification. In particular, DNA barcoding has been provided a great utility in biodiversity surveys. Moreover, the Conservation of biodiversity and natural plant resources is very important for the ecosystem (Kress & Horvitz 2005; Hosein *et al.* 2017). DNA barcoding has the potential to support species identification and discovery, vegetation, and floristic species surveys, in addition to studies on ecological forensics, all of which are critical to biodiversity management (Hollingsworth *et al.* 2016; Valentini *et al.* 2008; von Crautlein *et al.* 2011; Hosein *et. al.* 2017). Especially, DNA Barcoding and phylogenetic survey of the endemic plants are especially critical. DNA barcoding is a useful tool to quickly and accurately identify species. Also, DNA barcoding has the potential to prompt the discovery of new species.

Turkey has a significant and rich flora because of its geographical location and its paleogeography and ecological features. Turkey is the only country covered almost entirely by three of the world's 34 global biodiversity hotspots (the Caucasus, Iran-Anatolian, and the Mediterranean) (Sekercioglu *et al.* 2011). Mountain Erciyes is the highest mountain in Inner Anatolia. Moreover, mountain Erciyes in Turkey's Central Anatolia region due to rare species of plants in that area is one of Turkey's most important floristic area and has a different significance. (Byfield *et al.* 2003; Özhatay *et al.* 2009; Vural & Şapcı 2012). The Mount Erciyes has 89 families, 433 genera, and 1170 plant taxa (1116 species and 54 intraspecific taxa) (Vural & Aytaç 2005). While 44 of these taxa are endemic taxa from Turkey, of which 12 of them are known only from Mount Erciyes. One of these points endemic species is *Astragalus argaeus* Boiss. & Balansa and *A. stenosemioides* D.F. Chamb and V.A. Matthews (Vural & Aytaç 2005;

Vural & Şapcı 2012). In this study, the ability of DNA barcoding to confirm the identities of Critically Endangered endemic plant taxon (*Astragalus argaeus*) in Mountain Erciyes was assessed using a DNA barcode (*matK*).

Material And Methods

Taxon collection

Research materials were having gathered from Mount Erciyes, Kayseri, Turkey, and deposited at Herbarium of Erciyes University, Turkey (ERCH).

DNA isolation, amplification, and PCR product sequencing

Genomic DNA (gDNA) from leaf tissues was extracted by using a commercial DNA extraction kit (DNeasy Plant Mini Kit, Qiagen), following the manufacturer's protocol. Agarose gel electrophoresis of 1.5% was used to detect the presence of gDNA with ethidium bromide. The obtained genomic DNA was used in PCR for the amplification of the *matK* gene region using specific primer pairs 1R_KIM and 3F_KIM (Fazekas *et al.* 2008).

The PCR products were sequenced in both directions (Forward and Reverse) using an automatic sequencer (ABI 3100 Genetic Analyzer, RefGen Biotechnology, Ankara, Turkey). Reaction mixture used for the PCR amplifications contained a volume of 50 μl; 10 × Taq Buffer with $(NH4)2SO4$: 5 μl, 10 mM dNTP mix: 1.3 μl, 5u/ul Taq DNA polymerase (Thermo Scientific): 0.25 μl, 25 mM $MgCl2$: 3 μl, 10 mg/ml BSA: 3 μl, 10 uM each primer: 1 μl, and genomic DNA extract: 1 μl, dH2O: 34.45 μl). The following cycling parameters were used for *matK*, A pre-denaturation procedure consisting of 5 min at 95°C by 1 cycle, a denaturation step of 1 min at 94°C by 35 cycles, an annealing step of 1.30 min at 48°C by 35 cycles, an extension step of 1 min at 72°C by 35 cycles and an ending step of 10 min by 1 cycle.

Phylogenetic analyses

Sequences were aligned used for the multiple sequence alignment with Geneious Software 6R (available from http://www.geneious.com). Phylogenetic analyses were performed by using the sequences of *matK* gene regions obtained from GenBank (NCBI: The National Center for Biotechnology Information) and this study (Table 1). The related *matK* sequences of Astragalus species (23) were retrieved from the GenBank database. *Taraxacum alpinum* K.Koch. were used as the outgroup for the phylogenetic analysis. The best fit model for nucleotide substitution for sequences of *matK* gene regions was chosen using JModelTest 2.1 based on the corrected Akaike Information Criterion (AICc) and Bayesian Information Criterion (BIC) (Darriba *et al.* 2012). The BI analysis was performed by using BEAST v.1.8 (Drummond *et al.* 2012). All parameters were the effective sample sizes (ESS) > 200. Bayesian analysis was performed using four simultaneous runs of Metropolis-coupled Markov Chain Monte Carlo (MCMC) sampling for 10 million generations and under the uncorrelated lognormal relaxed clock model (Drummond *et al.* 2006) and also selected to Yule process for speciation as tree prior (Yule 1925). Checking of the Quality analysis was

done by using likelihood parameters and values, which were estimated by using different runs in Tracer V.1.5 (Rambaut & Drummond 2009).

Table 1
Accession numbers are available in GenBank database.

Taxon	Accession number
Astragalus boeticus	KX955075.1
Astragalus capito	KX955086.1
Astragalus caraganae	AB741312.1
Astragalus caspicus	MK958484.1
Astragalus cicer	KM387646.1
Astragalus commixtus	KX955090.1
Astragalus cornutus	KM387661.1
Astragalus corrugatus	AY920442.1
Astragalus dactylocarpus	AB741315.1
Astragalus daenensis	KX955095.1
Astragalus echinatus	JQ619978.1
Astragalus epiglottis	AB854563.1
Astragalus gaeobotrys	KX955113.1
Astragalus glycyphyllos	JN894914.1
Astragalus guttatus	KX955119.1
Astragalus hamosus	KX955121.1
Astragalus incertus	KX955127.1
Astragalus macrostachys	KX955143.1
Astragalus microcephalus	KX955147.1
Astragalus schizopterus	KM387635.1
Astragalus siliquosus	KX955182.1
Astragalus suberosus subsp. suberosus	KX955223.1
Astragalus tribuloides	KX955200.1
Astragalus argaeus	—
Astragalus stenosemioides	—

Results And Discussion

In this study, narrow endemic *A. argaeus* and *A. stenosemioides* were evaluated in the molecular data. Sequences of plastid DNA (*matK* gene) were successfully amplified from *A. argaeus* and *A. stenosemioides*. These sequences were analyzed together with sequences from GenBank and constructed phylogenetic tree. The phylogenetic tree was obtained by the Bayesian method using the BEAST program (Fig. 1). The phylogenetic tree constructed from *matK* sequences was clustered into the two main groups. *Astragalus caragnae* and *A. dactylocarpus* were grouped in the first lineage. Other species were classified in the second lineage. When we evaluated the DNA barcode-based tree, it shows that consistent with classical taxonomic classification. In this concept, the *matK* gene region sequences of the genus were useful to solve taxonomic problems. DNA barcoding involves sequencing a standard region of DNA as a tool for species identification. However, there has been no agreement on which region(s) should be used for barcoding land plants. An ideal DNA barcode should be routinely retrievable with a single primer pair. Additionally, it should be amenable to bidirectional sequencing with little requirement for manual editing of sequence traces, and also provide maximal discrimination among species. In this concept, there are several candidate barcode regions. One of these candidate barcode gene regions is the *matK* gene region (Ford *et al.* 2009). The *matK* gene region has a high rate of substitution, allowing substitutions to occur at almost the same rate in all triple codons (Hilu & Liang 1997, Hochbach *et al.* 2018). Besides, the excessive substitution rate of this gene is very effective in the formation of regions with parsimony information content and in obtaining strong phylogenetic predictions. Also, the rate of substitution of the *matK* gene provides an important guide in getting strong phylogenetic predictions and evaluating how the evolutionary past developed at different taxon levels (Johnson & Soltis 1994). However, some authors suggested that the *matK* gene region alone is insufficient for some plants (*Quercus* sp., *Salix* sp.) groups for determining DNA barcode. They stated that it is appropriate to use *matK* + *rbcL* gene regions together in such cases (Piredda *et al.* 2010; Ren *et al.* 2010; Von Crautlein *et al.* 2011). According to our study results, sequences of the *matK* gene region were enough variability for genus *Astragalus* ssp. Also, the *matK* gene region for DNA barcoding studies in genus *Astragalus* is a good candidate for DNA barcoding studies.

It is very important to protect and evaluate plant genetic resources. Hence, *A. argaeus* and *A. stenosemioides* are very important for endemic plants. These species have a very limited living area. Threat categories of *A. argaeus* and *A. stenosemioides* are Critically Endangered (CR). Especially, correct identification of endangered species is crucial for the future conservation of the species. The DNA barcoding is also proved valuable for accurate species identification as the important first step in conservation plans for threatened species (Kim *et al.* 2014). Barcoding studies can help on-going conservation prevention of the taxa in different ways. Species identification is the first important step in a correct evaluation of distribution, population abundance, and threats of target species (Hartvig *et al.* 2015). The present study was analyzed *matK* sequence data to determine if DNA barcoding could accurately identify genus *Astragalus*. Additionally, DNA sequences of endemic plants *A. argaeus* and *A. stenosemioides* were firstly determined in this study. The current study results show that *matK* gene region sequences can be used as potential a DNA barcoding region.

References

1. Byfield, A. Özhatay, N. & Atay, S. (2003). Türkiye'nin Önemli Bitki Alanları (Important Plant Areas in Turkey). WWF Türkiye (Doğal Hayatı Koruma Vakfı) Yayınları, İstanbul.

2. Carew, L., McMurtry, J. & Alster, F. (2005). Effects of lysine deficiencies on plasma levels of thyroid hormones, insulin-like growth factors I and II, liver and body weights, and feed intake in growing chickens. *Poult. Sci.* 84: 1045–1050.

3. Darriba, D., Taboada, G. L., Doallo, R. & Posada D. (2012). JModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* 9: 772. https://doi.org/10.1038/nmeth.2109

4. Drummond, A. J. Suchard, M. A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 22 (8): 1185–1192. https://doi.org/10.1093/molbev/mss075

5. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4 (5): 699–710. https://doi.org/10.1371/journal.pbio.0040088

6. Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster S. G., Husband B. C., Percy D. M., Hajibabaei M. & Barrett S. C. H. (2008). Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. *PLoS ONE* 3(7): e2802. https://doi.org/10.1371/journal.pone.0002802

7. Ford, C., Ayres, K., Toomey, N., Haider, N., Stahl, J., Kelly, L., Wikstrom, N., Hollingsworth, P, Duff, R., Hoot, S., Cowan, R., Chase, M. & Wilkinson, M. (2009). Selection of candidate coding DNA barcoding regions for use on land plants. J. *Linn. Soc., Bot.* 159(1): 1–11.

8. Hartvig, I., Czako, M., Kjær, E. D., Nielsen, L. R. & Theilade, I. (2015). The Use of DNA Barcoding in Identification and Conservation of Rosewood (*Dalbergia* spp.). *PLoS ONE* 10(9): e0138231. https://doi.org/10.1371/journal.pone.0138231.

9. Hilu, K. W. & Liang, H. (1997). The *matK* gene: sequence variation and application in plant systematics. *Am. J.Bot.* 84(6): 830–839.

10. Hochbach, A., Linder, H. P. & Röser, M. (2018). Nuclear genes, *matK* and the phylogeny of the Poales. *Taxon* 67(3): 521–536.

11. Hollingsworth, P. M., Li, D. Z., van der Bank, M. & Twyford, A. D. (2016). Telling plant species apart with DNA: From barcodes to genomes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 371: 20150338. https://doi.org/10.1098/rstb.2015.0338

12. Hosein, F. N., Austin, N., Maharaj, S., Johnson, W., Rostant, L., Ramdass, A. C. & Rampersad, S. N. (2017). Utility of DNA barcoding to identify rare endemic vascular plant species in Trinidad. *Ecol. Evol.* 7(18): 7311–7333.

13. Kim, H. M., Oh, S. H., Bhandari, G. S., Kim, C. S. & Park, C. W. (2014). DNA barcoding of Orchidaceae in Korea. *Mol.Ecol. Resour.* 14(3):499–507.

14. Özhatay, N., Kültür, Ş. & Aslan, S. (2009). Check-list of Additional Taxa to the Supplement Flora of Turkey IV. Turk. J. Bot. 33: 191–226.

15. Piredda, R., Simeone, M. C., Attimonelli, M., Bellarosa, R. & Schirone, B. (2010). Prospects of barcoding the Italian wild dendroflora: oaks reveal severe limitations to tracking species identity. *Mol.r Ecol. Resour.* 11(1): 72–83.

16. Rambaut, A. & Drummond, A. J. 2009. Tracer v 1.5. http://beast.bio.ed.ac.uk/Tracer.

17. Ren, B. Q., Xiang, X. G. & Chen, Z. D. 2010. Species identification of *Alnus* (Betulaceae) using nrDNA and cpDNA genetic markers. *Mol. Ecol.Resour.* 10(4): 594–605.

18. Sekercioglu, C., Anderson, S., Akcay, E., Bilgin, R., Can, O., Semiz, G., Tavşanoğlu, Ç., Soyumert, A., Yokes, M. B., Ipekdal, K., Sağlam, I., Yücel, M., & Dalfes, N. (2011). Turkey's globally important biodiversity in crisis. *Biol. Conser.* 144: 2752–2769. 10.1016/j.biocon.2011.06.025.

19. Valentini, A., Pompanon, F. & Taberlet, P. (2008). DNA barcoding for ecologist. *Trends Ecol. Evol.* 24: 110–117.

20. von Crautlein, M., Korpelainen, H., Pietiläinen, M. & Rikkinen, J. (2011). DNA barcoding: A tool for improved taxon identification and detection of species diversity. *Bio.Conserv.* 20: 373–389. https://doi.org/10.1007/s10531-010-9964-0.

21. Vural, C. & Şapcı, H. (2012). A single locality, endemic species, Astragalus argaeus (Fabaceae). Kew Bull. 67: 89–92.

22. Vural, C. & Aytaç, Z. (2005): The Flora of Erciyes Dağı (Kayseri, Turkey). Turk. J. Bot. 29: 185–236.

23. Yule, G. U. (1925). A Mathematical Theory of Evolution Based on the Conclusions of Dr. J.C. Willis, F.R.S. *J. R. Stat. Soc.* 213: 21–87. https://doi.org/10.2307/2341419
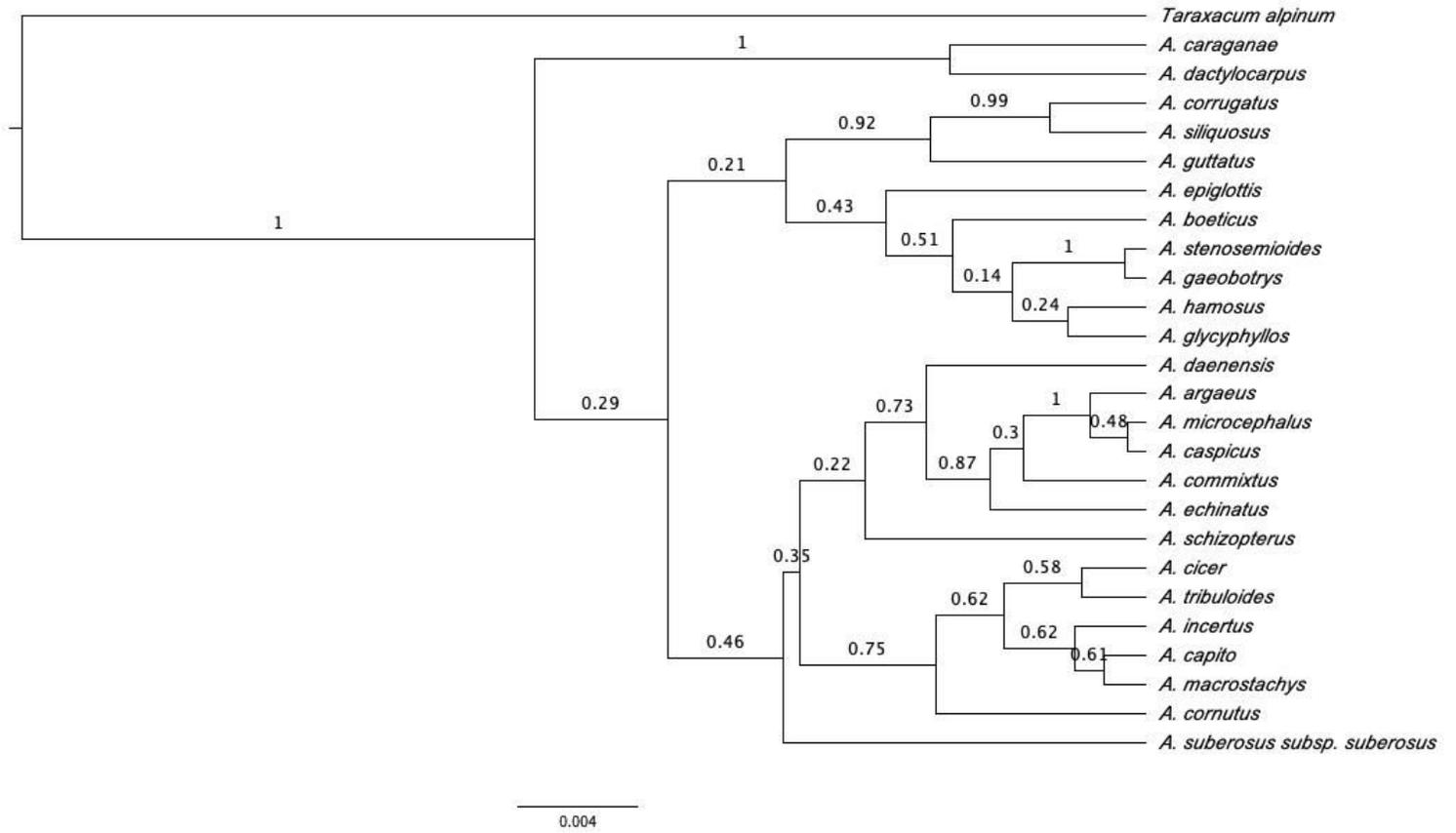
Figures

Figure 1

Phylogenetic tree using BEAST Bayesian inference based on cp DNA (matK gene region).