DNA barcoding and phylogenetic analysis of endemic *Astragalus nezaketiae* and *Vicia alpestris* subsp. *hypoleuca* (Fabaceae): Evidence from nrDNA ITS and cpDNA *matK* and *rbcL* sequences

Mehmet Y. PAKSOY¹, Emre SEVİNDİK²*, İsa BAŞKÖSE³

¹Munzur University, Department of Medical Services and Techniques, Tunceli Vocational School, 62000 Tunceli, Turkey; mypaksoy@gmail.com
²Aydın Adnan Menderes University, Faculty of Agriculture, Department of Agricultural Biotechnology, South Campus, Cakmar, Aydin, Turkey; ph.d-emre@hotmail.com (*corresponding author)
³Ankara University, Faculty of Science, Department of Biology, 06100 Ankara, Turkey; isabaskose@gmail.com

Abstract

In this study, we performed DNA barcoding and phylogenetic analysis using one nuclear (ITS) and two chloroplast DNA regions (*matK* and *rbcL*) of endemic *Astragalus nezaketiae* A. Duran & Aytaç and *Vicia alpestris* Stev. subsp. *hypoleuca* (Boiss.) Davis taxa in Turkey. PCR reactions were performed using universal primers. Sequences of the PCR products were edited using BioEdit and FinchTV software and contigs were obtained. All contigs were Blasted at NCBI and similarities were analysed. Using the MEGA 6.0 program, maximum likelihood trees were constructed including some sequences retrieved from NCBI. For *Astragalus nezaketiae*; in the ITS analysis, *Astragalus nezaketiae* appeared separately from other species, and for *matK*, *Astragalus nezaketiae* appeared together with *Astragalus cicer* L. However, *rbcL* tree was polytomic. For *Vicia alpestris* subsp. *hypoleuca*; in ITS, *rbcL* and *matK* results *Vicia alpestris* subsp. *hypoleuca* were found together with *Vicia cracca* L., *Vicia benghalensis* L. and *Vicia villosa* Roth species. Analysis of the combined data revealed similar results with all barcode regions for *Vicia alpestris* subsp. *hypoleuca* while different phylogenetic results were obtained for *Astragalus nezaketiae*.

*Keywords:* *Astragalus*; barcoding; ITS; *matK*; *rbcL*; *Vicia*

Introduction

Fabaceae (Leguminosae), the third largest family within the Angiosperms, includes 946 genera and over 24,500 accepted species (Manzione *et al.*, 2022). The Fabaceae, one of the herbaceous, shrub, woody and climbing plants with a cosmopolitan distribution, is one of the greatest flowering plant families in Turkey after Asteraceae with respect to number of species, and features 1013 species belonging to 71 genera (Toksoy *et al.*, 2015; İlcim and Behçet, 2016; Antonio-Domingues *et al.*, 2018). In addition, it consists of three subfamilies; Caesalpinioideae, Mimosoideae, and Papilionoideae (Han *et al.*, 2021a). Also known as the legume, pea or bean family, Fabaceae is an enormous and economically important member of flowering plants (Rahman and Parvin,
Additionally, ethnobotanical studies have reported that some taxa belonging to the Fabaceae family are used in folk medicine, along with providing tree resources and dyes, resins, insecticides, fibers, feed, and the like, from a socioeconomic point of view (Molares and Ladio, 2012; Zengin et al., 2015). *Astragalus* L. is the largest genus of vascular plants with nearly 2900 species. In Turkey, it is represented by 483 taxa which are divided into 63 sections (Karaman Erkul et al., 2022). *Astragalus* is widely distributed in the steppe environment of low or high mountains in the Irano-Turanian phytogeographic region of Turkey (Duran and Aytaç, 2005; Atasagun et al., 2021). Some species of *Astragalus* are used in farm and wild animal feed, foods, medicines, cosmetics, and as tea or as a source of herbal gum (Rios and Waterman, 1997). Furthermore, the medical use of the *Astragalus* species dates back ten centuries. In recent years, progress has been made in research on these species due to their anticancer compounds and medical uses (Ionkova et al., 2014). The genus *Vicia* L. belongs to the Fabaceae family and is the third largest family of flowering plants worldwide, containing about 210 species, with 22 divisions in two subgenera (*Vicia* and *Vicilla*), distributed throughout the temperate regions of Europe, Asia, and North and South America. In Turkey, 64 species, 22 subspecies, and 18 varieties of this genus have been recorded (Han et al., 2021b; Kaplan et al., 2021a, Kaplan et al., 2021b). The highest specific diversity of *Vicia* is found in Turkey and northwest Asia (Martin et al., 2018). The *Vicia* species, which are agriculturally and economically important, protein-rich, and valuable genetic resources, are used as high-quality green manure, medicine, food, cover, fodder, and ornamental and honey plants to reduce the occurrence of pests and diseases and increase soil fertility (Han et al., 2021b; Wu et al., 2021; Wu et al., 2020; Jo et al., 2022; Sun et al., 2022).

Plant molecular systematics has been developing rapidly, especially in the last 20 years. This development has started to contribute to molecular systematics with the use of DNA and amino acid sequence analysis and the discovery of new phylogenetic analysis methods (Inal and Karaca, 2019). The concept of "DNA barcode", which has been widely used in the literature, was first introduced by Hebert et al. (2003) and later managed to attract the attention of the science world (Filiz and Koç, 2012; Keskin and Atar, 2013). Thanks to this technique, it is easier to distinguish between species, and the identification process is quicker and more accurate (Koohdar et al., 2021). nrDNA ITS (internal transcribed spacer) and chloroplast DNA (*rbc*L, *atp*F-H, *rpo*B, *rpo*C1, *ndh*F, *mat*K, *trn*H-*npb*A, *rps*16-*trn*Q, *rpl*32-*trn*L and *trn*L-F) sequences are used frequently in molecular phylogenetic analyses of plants and barcoding (Yang et al., 2010; Filiz and Koç, 2012; Sun et al., 2012; Dong et al., 2012; Wang et al., 2013; Tchen et al., 2014; Lin et al., 2015; Dastpak et al., 2018; Kim et al., 2020; Inal and Karaca, 2019). In this study, we conducted DNA barcoding and phylogenetic analysis using one nuclear (ITS) and two chloroplast DNA regions (*mat*K and *rbc*L) of endemic *Astragalus nezaketiae* and *Vicia alpestris* subsp. *hypoleuca* taxa distributed in Turkey.

**Materials and Methods**

*Study area and genomic DNA isolation*

The study materials included the specimens of *Astragalus nezaketiae* and *Vicia alpestris* subsp. *hypoleuca* taxa. Taxa and the locations they are collected from are presented in Table 1. For genomic DNA isolation, a commercial kit (GeneMark Catalog No: DP022) was used.
Table 1. Location of endemic *Astragalus nezaketiae* and *Vicia alpestris* subsp. *hypoleuca* in Turkey

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Location and herbarium number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Astragalus nezaketiae</em></td>
<td>Tunceli province, Pülümür district, Bağırpaşa Mountain, calcareous rocky slopes, 2600 m a.s.l, 20 August 2019, Başköse and Paksoy 3004</td>
</tr>
<tr>
<td><em>Vicia alpestris</em> subsp. <em>hypoleuca</em></td>
<td>Tunceli province, Pülümür district, Bağırpaşa Mountain, calcareous rocky slopes, 2600 m a.s.l, 20 August 2019, Başköse and Paksoy 3005</td>
</tr>
</tbody>
</table>

**PCR amplifications, sequencing and phylogenetic analysis**

Primer sequences, PCR components and PCR protocols for amplifying the nrDNA ITS, cpDNA *matK* and *rbcL* regions via PCR are given in Table 2.

Table 2. Primers, PCR components and PCR protocols

<table>
<thead>
<tr>
<th>Primer name, 5'-3' sequences and references</th>
<th>PCR components</th>
<th>PCR Protocols</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITS5AF: 5'-CCTTATCATTTAGAGGAAGGAG-3' (White et al., 1990)</td>
<td>1 μL genomic DNA 1 μL primer (forward), 1 μL primer (reverse), 5 μL master mix (PCR buffer, 2 Mm MgCl₂, dNTP, 0.75 U Taq DNA polymerase) and 17 μL dH₂O</td>
<td>94°C, 5 min; 35× (94°C, 45 s., 50°C, 45 s., 72°C, 1 m.), 72°C, 10 min.</td>
</tr>
<tr>
<td>ITS4R: 5'-TCCTCCGCTTATTGATATGC-3' (White et al., 1990)</td>
<td>1 μL genomic DNA 1 μL primer (forward), 1 μL primer (reverse), 5 μL master mix (PCR buffer, 2 Mm MgCl₂, dNTP, 0.75 U Taq DNA polymerase) and 17 μL dH₂O</td>
<td>94°C, 4 min; 35× (94°C, 1 min), 48°C, 30 s; 72°C, 1 min; 72°C, 7 min.</td>
</tr>
<tr>
<td><em>matK</em>-472f: 5'-CCCRTYCATCTGAAATCTTGGTTC-3' (Yu et al., 2011)</td>
<td>1 μL genomic DNA 1 μL primer (forward), 1 μL primer (reverse), 5 μL master mix (PCR buffer, 2 Mm MgCl₂, dNTP, 0.75 U Taq DNA polymerase) and 17 μL dH₂O</td>
<td>94°C, 4 min; 35× (94°C, 1 min), 48°C, 30 s; 72°C, 1 min; 72°C, 7 min.</td>
</tr>
<tr>
<td><em>matK</em>-1248r: 5'-GCTTRATAATGAAAGATTTCCTGC-3' (Yu et al., 2011)</td>
<td>1 μL genomic DNA 1 μL primer (forward), 1 μL primer (reverse), 5 μL master mix (PCR buffer, 2 Mm MgCl₂, dNTP, 0.75 U Taq DNA polymerase) and 17 μL dH₂O</td>
<td>94°C, 4 min; 35× (94°C, 1 min), 48°C, 30 s; 72°C, 1 min; 72°C, 7 min.</td>
</tr>
<tr>
<td><em>rbcLa</em>-F: 5'-ATGTCACCACAAACAGAGACTAAAGC-3' (Levin, 2013)</td>
<td>1 μL genomic DNA 1 μL primer (forward), 1 μL primer (reverse), 5 μL master mix (PCR buffer, 2 Mm MgCl₂, dNTP, 0.75 U Taq DNA polymerase) and 17 μL dH₂O</td>
<td>95°C, 1 min; 35× (95°C, 30 s; 51°C, 30 s; 68°C, 1 min); 68°C, 5 min.</td>
</tr>
<tr>
<td><em>rbcLa</em>-R: 5'-GTAAAATCAAGTCCACCRCG-3' (Kress et al., 2009)</td>
<td>1 μL genomic DNA 1 μL primer (forward), 1 μL primer (reverse), 5 μL master mix (PCR buffer, 2 Mm MgCl₂, dNTP, 0.75 U Taq DNA polymerase) and 17 μL dH₂O</td>
<td>95°C, 1 min; 35× (95°C, 30 s; 51°C, 30 s; 68°C, 1 min); 68°C, 5 min.</td>
</tr>
</tbody>
</table>

PCR products were sent to the TRIOGEN (Istanbul/Turkey) biotechnology company for service procurement of their distillation and cycle sequencing reactions. The BioEdit 7.2.3 (Hall, 1999) and Finch TV version 1.4.0 programs were used in processing the DNA sequences which came as files in the ABI prism format. A contig was created using forward and reverse sequences related to each species. Specific bases misread by a device performing the sequencing reactions were visually corrected with the help of the Bioedit 7.2.3 and Finch TV version 1.4.0 programs based on the strength and cleanness of the signals (peaks) in the chromatogram. By this way, the contig sequences were obtained. The MEGA 6.0 program was used for aligning the sequences (Tamura et al., 2013). To evaluate the degree of support for clades, a bootstrap analysis (1,000 replicates) was applied (Felsenstein, 1985). In order to find out the phylogenetic relationships of the species whose sequences were obtained, MEGA 6.0 which is commonly used in biotechnology, molecular biology, genetic and bioinformatic research, was used. Data such as phylogenetic trees, genetic distance and nucleotide distance were acquired using the parameters and species in these programs.

**Results and Discussion**

The nrDNA ITS sequence of flowering plants is recommended as a core DNA barcode (Zhu et al., 2021) as it has high discriminatory power to distinguish closely related species. In plant phylogenetic studies of
closely related species and populations, this region is preferred by molecular taxonomists (Dizkirici and Koroglu, 2018). The ITS sequence length of the endemic *Astragalus nezaketiae* species was 671 bases, and the nucleotide composition was determined as 22.70% adenine, 25.50% cytosine, 27.90% guanine and 24.00% thymine. The maximum likelihood phylogenetic tree was constructed including ITS sequences of some *Astragalus* species obtained from NCBI. *Astragalus nezaketiae* species was separated from other species (Figure 1). The ITS sequence length of the endemic *Vicia alpestris* subsp. *hypoleuca* species was 680 bases, and the nucleotide composition was determined as 22.10% adenine, 24.10% cytosine, 27.40% guanine and 26.50% thymine. The maximum likelihood phylogenetic tree constructed including ITS sequences of some *Vicia* species from NCBI, consisted of two clades (Figure 2). *Vicia alpestris* subsp. *hypoleuca* species coexisted with *Vicia cracca*, *Vicia benghalensis* and *Vicia villosa* in subclade A, and this branch received a bootstrap value of 94%.

---

**Figure 1.** The maximum likelihood tree generated using nrDNA ITS sequences of *Astragalus nezaketiae* and other *Astragalus* sequences retrieved from NCBI

**Figure 2.** The maximum likelihood tree generated using nrDNA ITS sequences of *Vicia alpestris* subsp. *hypoleuca* and other *Vicia* sequences retrieved from NCBI
Advances in DNA sequencing techniques have allowed the studies of phylogenetic relationships with the widespread use of short DNA fragments especially of the chloroplast genome (Penjor et al., 2010). DNA barcoding relies on information encoded in the nucleotide sequences of a genome’s standard region as a tool for species identification. The CBOL working group proposed Rubisco large subunit (rbcL) and matK as the standard plant barcode based on sequence quality and species separation levels (CBOL Plant Working Group, 2009; Bafeel et al., 2011). The matK sequence length of the Astragalus nezaketiae was 738 bases, and the nucleotide composition was determined as adenine 30.10%, cytosine 15.90%, guanine 15.40% and thymine 38.60%. Including matK sequences of some Astragalus species from NCBI, a maximum likelihood phylogenetic tree composed of two clades was constructed (Figure 3). Astragalus nezaketiae and Astragalus cicer were placed in clade 2 which was supported with a 62% bootstrap value. The rbcL sequence length of Astragalus nezaketiae was 546 bases, and the nucleotide composition was determined as adenine 26.70%, cytosine 20.10%, guanine 23.30% and thymine 29.90%. In a maximum likelihood phylogenetic tree constructed including rbcL sequences of some Astragalus species from NCBI, Astragalus nezaketiae was found to be polytomic (Figure 4).

![Figure 3](image-url)  
Figure 3. The maximum likelihood tree generated using cpDNA matK sequences of Astragalus nezaketiae and other Astragalus sequences retrieved from NCBI

![Figure 4](image-url)  
Figure 4. The maximum likelihood tree generated using cpDNA rbcL sequences of Astragalus nezaketiae and other Astragalus sequences retrieved from NCBI
The *matK* sequence length of the *Vicia alpestris* subsp. *hypoleuca* species was 739 bases, and the nucleotide composition was determined as 38.84% adenine, 15.43% cytosine, 16.24% guanine and 29.50% thymine. The maximum likelihood phylogenetic tree constructed including *matK* sequences of some *Vicia* species from NCBI consisted of two clades (Figure 5). *Vicia alpestris* subsp. *hypoleuca* species was placed with *Vicia alpestris*, *Vicia cracca*, *Vicia benghalensis* and *Vicia villosa* species in group A, and this group received a bootstrap value of 97%. The *rbcL* sequence length of *Vicia alpestris* subsp. *hypoleuca* was 546 bases, and the nucleotide composition included adenine 28.2%, cytosine 20.1%, guanine 22% and thymine 29.7%. In the maximum likelihood phylogenetic tree constructed including the *rbcL* sequences of some *Vicia* species from NCBI, *Vicia alpestris* subsp. *hypoleuca* found in the same group with *Vicia cracca*, *Vicia benghalensis* and *Vicia villosa*, and this group received a 60% bootstrap value (Figure 6).

**Figure 5.** The maximum likelihood tree generated using cpDNA *matK* sequences of *Vicia alpestris* subsp. *hypoleuca* and other *Vicia* sequences retrieved from NCBI

**Figure 6.** The maximum likelihood tree generated using cpDNA *rbcL* sequences of *Vicia alpestris* subsp. *hypoleuca* and other *Vicia* sequences retrieved from NCBI
Dizkirici et al. (2014) in ITS analysis study have determined Astragalus nezaketiae together with Astragalus humillimus Barneby, Astragalus sanguinolentus M. Bieb., Astragalus glaucocephalus Bunge, Astragalus achundovii Grossh. and Astragalus schizopterus Boiss. species. A cpDNA matK analysis study of Tekpinar et al. (2016) reported Astragalus nezaketiae together with Astragalus czorochensis Kharadze, Astragalus humillimus, Astragalus sanguinolentus and Astragalus schizopterus species. Kaplan et al. (2021a) in their cpDNA trnL intron UPGMA analysis determined Vicia sativa subsp. sativa, Vicia sativa subsp. nigra var. nigra, Vicia sativa subsp. amphicarpa, Vicia lathyroides L. in a group (bootstrap 100%), while Vicia lutea var. hirta, Vicia villosa subsp. varia, Vicia villosa subsp. eriocarpa and Vicia cracca subsp. stenophylla within a group identified In our ITS, rbcL and matK analysis results, the Vicia alpestris subsp. hypoleuca, Vicia cracca, Vicia benghalensis and Vicia villosa were together determined. Kaplan et al. (2021b) in their ITS analysis found that the Vicia sativa subsp. sativa, Vicia sativa subsp. amphicarpa, Vicia sativa subsp. nigra and Vicia lathyroides were together, whereas the Vicia villosa subsp. varia, Vicia villosa subsp. eriocarpa and Vicia cracca subsp. stenophylla and Vicia lutea var. hirta were in a different group. In the NJ phylogenetic tree generated using ITS2 + matK + psbA-trnH + rbcL sequences, Raveendar et al. (2017) determined the Vicia sativa and the Vicia hirsuta (L.) Grey within a group (bootstrap 100%), Vicia tetrasperma and Vicia articulate within a group (bootstrap 100%), the Vicia lutea L. and Vicia peregrina L. within another group (bootstrap 99%) and Vicia cracca, Vicia benghalensis, and Vicia villosa in another group (bootstrap 94%). Bozkurt et al. (2013) in their ISSR analysis determined V. sativa subsp. sativa, V. sativa subsp. nigra var. nigra, V. sativa subsp incisa (M. Bieb.) Arcang. var. cordata and V. peregrina and V. hybrida species within a group and the V. cracca subspecies together with the V. palaestina Boiss. Endo et al. (2010), in an ITS analysis performed with the Bayesian tree method determined the Vicia cracca and Vicia benghalensis species (bootstrap 100%) within a group and the Vicia sativa var. angustifolia, Vicia sativa and Vicia americana Muhl. ex Willd within another group (bootstrap 99%). In the cpDNA matK analysis (Bayesian tree), they determined the Vicia hirsuta and Vicia articulata species within a group (bootstrap 100%); Vicia lutea, Vicia sativa and Vicia americana together, and the Vicia villosa and Vicia benghalensis species within a separate group (bootstrap 100%).

Conclusions

In this report, we studied endemic Astragalus nezaketiae and Vicia alpestris subsp. hypoleuca species distributed in Turkey to run their phylogenetic analysis using ITS, matK and rbcL sequences as barcode regions. Of these barcode regions, only the rbcL region did not yield efficient results for Astragalus nezaketiae. For Vicia alpestris subsp. hypoleuca, the phylogenetic analysis of the ITS, matK and rbcL barcode regions were congruent. We believe the result of this study will guide the further phylogenetic analyses of Astragalus nezaketiae and Vicia alpestris subsp. hypoleuca species.

Authors’ Contributions

M.Y.P and İ.B collected plants samples. The experiments were performed and analyzed by E.S. E.S wrote the paper. All authors read and approved the final manuscript.
Ethical approval (for researches involving animals or humans)

Not applicable.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

References


The journal offers free, immediate, and unrestricted access to peer-reviewed research and scholarly work. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles, or use them for any other lawful purpose, without asking prior permission from the publisher or the author.

**License** - Articles published in *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* are Open-Access, distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) License.

© Articles by the authors; Licensee UASVM and SHST, Cluj-Napoca, Romania. The journal allows the author(s) to hold the copyright/to retain publishing rights without restriction.

**Notes:**
- **Material disclaimer:** The authors are fully responsible for their work and they hold sole responsibility for the articles published in the journal.
- **Maps and affiliations:** The publisher stay neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- **Responsibilities:** The editors, editorial board and publisher do not assume any responsibility for the article’s contents and for the authors’ views expressed in their contributions. The statements and opinions published represent the views of the authors or persons to whom they are credited. Publication of research information does not constitute a recommendation or endorsement of products involved.