Internal Transcribed Spacer (ITS) Fails Barcoding of the Genus *Neotinea* Rchb.f. (Orchidaceae)

Kaan HÜRKAN\textsuperscript{a}*, Kemal Melih TAŞKIN\textsuperscript{b}

\textsuperscript{a}İğdır University, Faculty of Agriculture, Department of Agricultural Biotechnology, İğdır, TURKEY
\textsuperscript{b}Çanakkale Onsekiz Mart University, Faculty of Sciences and Arts, Department of Molecular Biology and Genetics, Çanakkale, TURKEY

**ARTICLE INFO**
Research Article
Corresponding Author: Kaan HÜRKAN, E-mail: kaanhurkan@hotmail.com
Received: 05 September 2019 / Revised: 19 October 2019 / Accepted: 29 October 2019 / Online: 18 January 2021

**ABSTRACT**
Internal Transcribed Spacer (ITS) is one of the most used barcoding regions for the molecular phylogenetics and barcoding of orchids. Our aim in this study is to test the reliability of ITS on barcoding of closely related *Neotinea* spp., including *Neotinea tridentata*, *Neotinea ustulata* subsp. *ustulata* and *Neotinea ustulata* subsp. *aestivalis*, by comparing it to the accD-psaI intergenic spacer of the plastid DNA. Both ITS and accD-psaI regions were amplified by specific primer sets and sequenced. Phylogenetic trees were regenerated by using Maximum Parsimony approach. The results showed that ITS separated some *N. tridentata* samples of Turkish, Greek, Hungarian and Croatian samples from the others on the phylogenetic trees due to the incomplete lineage sorting. In contrast to ITS, the accD-psaI marker could successfully separate *N. tridentata* and *N. ustulata* samples according to a priori species classification. Our findings refer to a hybridisation story between some *N. tridentata* and *N. ustulata*. We propose not to use ITS sequences directly as a barcode and to reconstruct the phylogeny of the *Neotinea* group. Instead, the inclusion of other nuclear regions such as LFY, ADH, etc., or utilisation of whole genome sequencing could give better barcoding results.

Keywords: Orchids, accD-psaI, DNA barcoding, phylogenetic incongruence.

© Ankara University, Faculty of Agriculture

1. Introduction

The orchid family (Orchidaceae) is the second largest flowering plant family represented by 899 genera and 27801 species (The Plant List 2013). Therefore, it is necessary to develop a reliable molecular identification method. Since the 1990s, using molecular markers on the phylogenetic studies of orchids have been increasing incrementally (Chase et al. 2000; Bateman et al. 2003; Sramko et al. 2014). Plant taxonomists have used short sequences of both ribosomal deoxyribonucleic acid (rDNA) and chloroplast deoxyribonucleic acid (cpDNA) as molecular markers to identify species and to reconstruct phylogeny (Zimmer & Wen 2012). Rapidly evolving DNA regions are needed for species-level barcoding (Sramkó et al. 2011; Zimmer & Wen 2012). The Internal Transcribed Spacer of the nuclear ribosomal 18S-5.8S-26S cistron (ITS) is one of the most extensively used molecular markers in orchid identification and molecular phylogenetic since the 1990s (Baldwin et al. 1995; Pridgeon et al. 1997; Bateman et al. 2005; Nieto-Feliner & Rosselló 2007; Hollingsworth 2008; Sramko et al. 2014; Li et al. 2015). The main advantages of ITS are (1) universality of primer sets (White et al. 1990), (2) ease of amplification even from historic specimens (i.e., herbarium specimens) due to multi-copy characteristic, and (3) favourable size of the region (~700 bp) (Baldwin et al. 1995). Some researchers point out the concerted evolution of ITS region (Bailey 2003; Gulyás et al. 2005; Pillon et al. 2007; Nieto-Feliner & Rosselló 2007). Concerted evolution is a process that homogenizes the rDNA array within an organism and reduces intra-individual variation (Liao 2008). accD-psaI barcoding region covers partial acetyl-CoA carboxylase enzyme (accD) and Photosystem I (psaI) coding regions and an intergenic region on the chloroplast genome.

In this study, we would like to test the utilities of ITS and accD-psaI regions for DNA barcoding of closely related *Neotinea* (Rchb.f.) taxa; *Neotinea tridentata* (Scop.) R. M. Bateman, Pridgeon & M. W. Chase, *Neotinea ustulata* subsp. *ustulata* (L.) R. M. Bateman, Pridgeon & M. W. Chase, *Neotinea ustulata* subsp. *aestivalis* (Kümpel) Kolnık, Vlčko & Ditě by comparing the topology of the phylogenetic trees.
2. Material and Methods

2.1. Plant material and DNA extraction

The field-collected leaf samples placed in silica-gel sachets and stored at room temperature until DNA extraction. The samplings of *N. tridentata*, *N. ustulata* subsp. *ustulata*, *N. ustulata* subsp. *aestivalis*, *Neotinea maculata* (Desf.) Stearn and *Neotinea conica* (Wild.) R. M. Bateman, were obtained from Mediterranean, west sub-Mediterranean, middle sub-Mediterranean, Balkan sub-Mediterranean and east sub-Mediterranean regions (Figure 1: 1A and 1B). Altogether, we analysed 23 samples of *Neotinea* spp. from the field and 9 DNA sequences from the nucleotide collection GenBank (Supplementary Data 1). We chose *Ophrys phrygia* H. Fleischm. & Bornm. (Genbank accession no: MH050859) and *Platanthera dilatata* (Pursh) Lindl. ex L. C. Beck (Genbank accession no: JX484921) as out-groups.

Figure 1- 1A; Sampling map. A: Mediterranean, B; West sub-Mediterranean, C; Mid sub-Mediterranean, D; Balkan sub-Mediterranean and E; East sub-Mediterranean. *(Platanthera dilatata was excluded).* 1B; Images of *Neotinea maculata* (A), *Neotinea tridentata* (B) and *Neotinea ustulata* (C). The image of *N. ustulata* (Vladan Djordjevic).

The DNA extraction was performed according to the cetyltrimethylammonium bromide (CTAB) protocol (Doyle & Doyle 1990). Approximately 10 mg of dried leaves was used for the extraction.

2.2. Choice of molecular markers for molecular barcoding

We used ITS as a nuclear-molecular marker to test its molecular barcoding utilities for *Neotinea* spp. which were collected from a wide geographic range. We also added the plastid-encoded *accD-psaI* intergenic region to be compared with ITS. Thus, we would be able to check the hybridisation origin of *N. tridentata*, as well.

2.3. PCR amplification, sequencing and phylogenetic reconstruction

We used angiosperm-specific ITS1A (Gulyás et al. 2005) and universal ITS4 (White et al. 1990) primer pair to amplify the whole ITS region. 25 μL of the PCR reaction mixture consisted of 0.1 volume of 10x High-Fidelity Buffer (Fermentas, USA), 0.5 μL 10 mM of each dNTP (Fermentas, USA), 2 μL 25 mM MgCl₂ (Fermentas, USA), 1.25 μL 20 mg mL⁻¹ BSA, 0.5 μL 10 μM primers (Thermo Scientific, USA), 0.1 μL 5 Unit High-Fidelity polymerase (Fermentas, USA), 16.65 μL water and 1μL (~5 ng μL⁻¹) genomic DNA. PCR amplification of ITS was performed on BIO-RAD PTC-200 thermal cycler using the following PCR profile: 94 °C for 2 min first denaturation, followed by 33 cycles of denaturation for 20 s at 94 °C, annealing for 30 s at 51 °C and extension for 1 min at 72 °C and finalized by final extension for 10 min at 72 °C. PCR products were then sent to Macrogen Inc. (The Netherlands) for Sanger sequencing (Applied Biosystems 3100 Genetic Analyzer) using ITS1A primer. Double signal peaks and Additive polymorphic sites (APS) were coded according to the International Union of Pure and Applied Chemistry (IUPAC) ambiguity symbols. We annotated the whole ITS region into the parts ITS1, 5.8s and ITS2 according to the sequences retrieved from the GenBank.

We used accD-f and psaI-r primer pair to amplify the *accD-psaI* region (Small et al. 1998). The PCR mixture was the same used for ITS amplification. PCR profile was applied as follows; 94 °C for 3 min first denaturation, followed by 34 cycles of denaturation for 30 s at 94 °C, annealing for 30 s at 55 °C and extension for 1 min at 72 °C and finalized by a final extension for 10 min at 72 °C. Amplified samples were sequenced in both directions using accD-f and psaI-r primers.
For the ITS and accD-psaI regions, the sequences were aligned by ClustalW (Larkin et al. 2007) under the bioinformatics software package Geneious R8 (Kearse et al. 2012). We used Maximum Parsimony (MP) heuristic search by PAUP 4.0a (build 159) (Swofford 2003). The MP-heuristic-search settings were: trees obtained from 1000 random replicates holding one tree at each step for stepwise addition, gaps are treated as missing, out-group: O. phrygia (ITS) and P. dilatata (accD-psaI), steepest descent option not in effect, and tree-bisection-reconnection (TBR) was used as branch swapping.

3. Results

3.1. Information obtained from DNA markers

ITS and accD-psaI regions were successfully amplified and sent to the sequencing service. The sequence files were imported into Geneious R8 software and checked for quality values. We excluded three samples since sequencing qualities of two samples from Italy, N. tridentata from Mt. Gargano and Toskana for ITS and accD-psaI, and one sample from Hungary, N. tridentata from Tokaj, for ITS were unacceptable for both ITS regions. We detected three interspecific length polymorphisms among five studied taxa; N. conica has 613 bp ITS, N. maculata has 609 bp ITS length. N. tridentata, N. ustulata subsp. ustulata and N. ustulata subsp. aestivalis have identical 616 bp ITS length. Whereas those polymorphisms are observed on ITS1 (ranged 234-243 bp) and ITS2 (ranged 217-222 bp) parts, there was no length polymorphism on the coding 5.8S part (Table 1). There was no length polymorphism on the accD-psaI region.

Table 1- ITS length polymorphisms on Neotinea spp

| Sample       | Accession No | Collecting site       | Lengths in base-pair |
|--------------|--------------|-----------------------|----------------------|
|              |              |                       | ITS1 | 5.8S | ITS2 | Total Length |
| N. conica    | AY364880     | Caceres-Spain          | 243  | 153  | 217  | 613          |
| N. maculata  | AM711744     | Kythera-Greece         | 234  | 153  | 222  | 609          |
| N. maculata  | MH050840     | Likouria-Greece        | 234  | 153  | 222  | 609          |
| N. maculata  | AY364873     | Tras os Montes-Portugal| 234  | 153  | 222  | 609          |
| All N. tridentata samples | All Samples | All Samples           | 242  | 153  | 221  | 616          |

Table 2- Initial comparison of DNA markers used in the study. O. phrygia (ITS) and P. dilatata (accD-psaI) are treated as out-groups relative to in-group samples

| DNA Region | Origin | Length range | Number of in-group polymorphic sites | Parsimony informative characters | Variability |
|------------|--------|--------------|-------------------------------------|---------------------------------|-------------|
|            |        |              | ITS1 | 5.8S | ITS2 | Aligned length | Out-group | In-group | Out-group | In-group |
| ITS        | Nucleus| 609-623      | 8    | 0    | 3    | 37  | 632 | 111 | 44 | 17.56% | 7.0% |
| accD-psaI  | Chloroplast| 923         | -    | -    | -    | 8   | 933 | 8   | 3  | 0.86%  | 0.32% |

Accordingly, the sequence variability of each DNA marker (Table 2) ITS region outperforms the chloroplast region accD-psaI in several variable positions (44 to 3 in-group), percentage within the aligned length (7.0% to 0.32% in-group) and parsimony informative characters (37 to 8). This comparison also shows that the ITS1 part of the ITS region has more variable sites (3.3% to 0.9%) than the ITS2 part. The aligned length was calculated as 632 bp for the ITS region and 933 bp for accD-psaI region. The protein coding 5.8S part has no variable sites on the studied samples, as expected. We observed APS on the chromatogram data of ITS sequences caused by the multi-copy nature and incomplete concerted evolution feature of the region (Table 3 and Figure 2).
3.2. Phylogenetic reconstruction

The MP phylogenetic analysis of the ITS region used 37 parsimony-informative characters and retained one most-parsimonious tree in 8028 rearrangement trials. The score of the best tree was 119. The MP tree showed moderate to high bootstrap support values (75 to 100).

On the ITS MP tree, *O. phrygia* was designated as the out-group and main clades consisted of *N. maculata* and *N. tridentata/N. ustulata* block. *N. tridentata* samples from Turkey (Adana, Antalya and Trabzon), Hungary (Siklos) and Croatia (Ucka) were apart from the main *N. tridentata/N. ustulata* clade. ITS sequences of those samples showed differences from other *N. tridentata* samples on the alignment as well since having APS. *N. tridentata* samples collected from Adana, Antalya and

---

Table 3- Additive Polymorphic Sites (APS) on ITS sequences

| Sample                          | Number of APS | Position (bp) | Nucleotides (IUPAC Symbols) |
|---------------------------------|---------------|---------------|-----------------------------|
| *N. tridentata* (Bahcesir-TR)   | 1             | 432           | G or A (R)                  |
| *N. tridentata* (Canakkale-TR)  | 1             | 513           | G or A (R)                  |
| *N. ustulata* subsp. aestivalis (Kiralyko-RO) | 2         | 513, 606      | G or A (R)                  |
| *N. ustulata* (Velka nad Velickou-CZ) | 3     | 513, 514, 606 | G or A (R)                  |
| *N. tridentata* (Adana-TR)      | 1             | 555           | G or A (R)                  |
| *N. tridentata* (Trabzon-TR)    | 1             | 555           | G or A (R)                  |
| *N. tridentata* (Antalya-TR)    | 1             | 555           | G or A (R)                  |

---

**Figure 2- Additive Polymorphic Sites (APS) of ITS sequences**
Trabzon samples have APS on the 559th position of the alignment (555th position on the sequence file), and on those positions, the nucleotides were either Adenine or Guanine. *N. tridentata* samples from Hungary (Siklos) and Croatia (Ucka) have different nucleotide (Thymine instead of Cytosine) on the 593th position of the alignment (589th position on the sequence file). *N. tridentata* sample from Greece (Pelion) had one nucleotide difference (Cytosine instead Thymine) on the 229th position of the alignment (226th position on the sequence file) than other *N. tridentata* samples. There was no APS on the sequence of those samples.

According to the ITS MP tree, *N. tridentata* samples collected from Adana, Trabzon, Antalya (Turkey), Siklos (Hungary), Ucka (Croatia) and Pelion (Greece) are sister to other *N. tridentata* samples. This results in polytomy among *N. tridentata*.

The MP-phylogenetic analysis of the accD-psal region used 8 parsimony-informative characters and retained one most-parsimonious tree in 28776 rearrangement trials. The score of the best tree was 381. On the phylogram of accD-psal sequences, *N. ustulata* samples placed together on one clade did not separate from each other as on the ITS tree (Figure 3). Three *N. ustulata* samples from Bulgaria, Czech Republic and Romania were separated on the tree, and this separation was highly supported by bootstrap. However, *N. ustulata* subsp. *ustulata* from Hungary was grouped with the *N. tridentata* samples. *N. tridentata* showed polytomy on the tree and, therefore, *N. ustulata* and *N. tridentata* samples were placed on the same branch.

4. Discussion

ITS region of nuclear ribosomal DNA is the most popular region for molecular phylogenetic studies for diverse plant groups including orchids, although it has some conspicuous drawbacks such as ITS due to incomplete concerted evolution (Nieto-Feliner 2003). ITS are observed when two nucleotides involved in a polymorphic site and those are very common when a hybridisation event on the background (Fuertes Aguilar & Nieto-Feliner 2003). Concerted evolution concept described as a genetic process by which repetitive DNA sequences are homogenized among the genome of different species, and it is known that the ITS has concerted evolution characteristic (Liao 2008). APS are indicative of incomplete concerted evolution due to a hybridisation event has not been homogenized within the genome (Bailey 2003). Those polymorphic ITS copies have been reported by authors in the literature (Mayol & Rosselló 2001, Won & Renner 2005). Our sequencing results revealed that the ITS has APS on the sequence chromatogram data (Figure 2). This finding refers to a hybridisation event and incomplete concerted evolution on *Neotinea* spp. particularly distributed in Asia Minor. These circumstances had an impact on the resolution power and caused sample misplacements on the phylogenetic trees. Additionally, incomplete concerted evolution of ITS marker on *Neotinea* spp. is in contradiction with the DNA barcode concept, since double nucleotide peaks are common throughout the sequence.

Our two phylogenetic trees have provided contrasting results for closely related *Neotinea* spp. On the ITS tree, *N. tridentata* samples from Adana, Trabzon, Antalya (Turkey), Siklos (Hungary), Ucka (Croatia) and Pelion (Greece) were separated from other *N. tridentata* samples due to having APS which comes from different copies of the ITS sequence in the genome. On the phylogram (Figure 3) it is also seen that *N. tridentata* samples from Adana, Trabzon and Antalya separated from other samples collected from Turkey. The issue with ITS region was discussed extensively in the study of Nieto-Feliner & Rosselló (2007). In the study, the authors listed the drawbacks of the ITS region and stressed its multi-copy nature which causes less confident results on phylogenetic studies. On the other study on which ITS was used, neither direct sequencing nor cloning efforts provided sufficient data for the phylogeny of *Himantoglossum* W.D.J. Koch s.l. (Orchidaceae) (Sramkó et al. 2011). In our study, in contrast to the study of (Pridgeon et al. 1997), the resolution power of the ITS region was insufficient to separate closely related *N. tridentata* and *N. ustulata* species that we sampled.

Our second molecular phylogenetic tool, accD-psal region, performed better than ITS for the topology of the phylogenetic tree. On the tree, *N. tridentata* and *N. ustulata* species were separated except one *N. ustulata* sample from Varpalota (Hungary). When we compared the topologies of our two phylogenetic trees side-by-side, the incongruence is clear. While accD-psal separate *N. tridentata* and *N. ustulata*, ITS does not. The contrasting incongruences on the two differently inherited molecular markers, i.e. cpDNA and ITS, refer to a hybridisation (Rieseberg et al. 1996; Gulyás et al. 2005; Sramkó et al. 2008, Kim & Donoghue 2008, Sramkó et al. 2014). In our results, the clear incongruence refers to the complex history of *Neotinea* genus and possible hybridisation between *N. tridentata* and *N. ustulata*, also. ITS region failed at the genus *Neotinea* since incomplete concerted evolution and hybridization background causes double-peaks and additive polymorphic sites on the sequence.

According to the statistics, in the 66% of papers published between 1998 and 2002, the ITS was used and 34% of them used ITS as the only molecular marker (Alvarez & Wendel 2003). This statistic shows the popularity of the ITS marker in the beginning and the improvement stage of the molecular phylogenetic studies. Later, after 2005 to date, the rate of ITS usage as the only molecular marker was decreased to 15.7%, and other molecular markers, i.e. mitochondrial or chloroplast sequences, were included to support the data gathered from the ITS (Nieto-Feliner & Rosselló 2007) that indicated the doubts on the ITS and attempts to find new molecular markers.
Figure 3- A: MP Phylogram of Neotinea spp. based on ITS sequences. The regions of origin are shown in parenthesis as country codes. O. phrygia is selected as the out-group. B: MP Cladogram of Neotinea spp. based on accD-psaI sequence. The regions of origin are shown in parenthesis as country codes. Bootstrap-support values are shown on nodes. P. dilatata is selected as the out-group. The pie chart represents the regions where differentiated samples collected from

5. Conclusions

On the history of molecular markers, the ITS region is still popular. It is not possible to abandon ITS as a molecular marker from the recent phylogenetic studies, particularly on orchids. However, researchers who study closely related species such as Neotinea spp., must consider the drawback of the marker. The strong conflicts between ITS and accD-psaI point out a possible hybridisation story. Fortunately, there are new technological ways of securing correct reconstruction of molecular phylogenetic trees and barcoding the closely related species, e.g. utilization of Next Generation Sequencing (NGS) techniques like RAD-seq.

Acknowledgements

We are grateful to Dr Gábor Sramkó for sharing his Neotinea spp. collection and his contributions improving the manuscript. We thank Cathy Seither for language proofing. This study was financially supported by Çanakkale Onsekiz Mart University, Scientific Research Coordination Unit; Project number: BAP FBA-2015-415. This study is based on the Doctoral Thesis titled “Studies on Phylogenetic Analysis of Neotinea ustulata and Neotinea tridentata (Orchidaceae)” written by Kaan Hürkan.

Supplementary Table

The whole list of the specimens sampled is available on Table S1 with the GenBank-accession numbers.

References

Alvarez I & Wendel J F (2003). Ribosomal ITS sequences and plant phylogenetic inference. Molecular Phylogenetics and Evolution 29(3): 417-434 doi: 10.1016/S1055-7903(03)00208-2
**Supplementary Table S1.** Table of Sampling and GenBank accession numbers*

| Sample               | GenBank Accession Number | Collecting site          |
|----------------------|--------------------------|--------------------------|
|                      | nrITS | acc-D-psa1 |                      |
| N. conica            | AY364880         | -                        | Caceres-Spain          |
| N. maculata          | MH050840 | MH400962 | Kythera-Greece       |
| N. maculata          | AY364873         | -                        | Tras os Montes-Portugal |
| N. tridentata        | MH050841 | MH400963 | Adana-Turkey        |
| N. tridentata        | MH050842 | MH400964 | Antalya-Turkey      |
| N. tridentata        | MH050843 | MH400965 | Bâlcesr-Turkey       |
| N. tridentata        | MH050844 | MH400966 | Çanakkale-Turkey     |
| N. tridentata        | MH050845 | MH400967 | Edirne-Turkey        |
| N. tridentata        | MH050846 | MH400968 | Gélibolu-Turkey      |
| N. tridentata        | MH050847         | -                        | Hatay-Turkey          |
| N. tridentata        | MH050848 | MH400969 | Kahramanmaraş-Turkey |
| N. tridentata        | MH050849 | MH400970 | Kastamonu-Turkey     |
| N. tridentata        | MH050850 | MH400971 | Litochoron-Greece    |
| N. ustulata subsp. ustulata | MH050858 | MH400980 | Lyalintsi-Bulgaria  |
| N. tridentata        | MH050851 | MH400972 | Osmaniye-Turkey      |
| N. tridentata        | KF499509         | -                        | Pelion-Greece         |
| N. tridentata        | MH050852 | MH400973 | Sîklos-Hungary       |
| N. tridentata        | MH050853 | MH400974 | Tekirdağ-Turkey      |
| N. tridentata        | MH050854 | MH400976 | Trabzon-Turkey       |
| N. tridentata        | MH400975         | -                        | Tokaj-Hungary         |
| N. tridentata        | MH050855 | MH400977 | Uçka-Croatia         |
| N. ustulata subsp. ustulata | FR750397 | -                        | Budapest-Hungary      |
| N. ustulata subsp. ustulata | AY364883 | -                        | Leon-Spain            |
| N. ustulata subsp. ustulata | AY364883 | -                        | Salamaca-Spain        |
| N. ustulata subsp. ustulata | AY014549 | -                        | Valais-Switzerland    |
| N. ustulata subsp. aestivalis | MH050857 | MH400979 | Varpalota-Hungary    |
| N. ustulata subsp. aestivalis | MH053407 | MH400981 | Velka nad Velickou-Czechia |
| N. ustulata subsp. aestivalis | MH050856 | MH400978 | Kiralyko-Romania     |
| Ophrys phrygia       | MH050859         | -                        | Antalya-Turkey        |
| Platanthera dilatata | - | JX484921 | -                      |
| Neotinea tridentata  | - | - | Toskana-Italy        |
| Neotinea tridentata  | - | - | Monte Gargano-Italy  |

*The dashes represents no reliable sequencing quality or data not available on GenBank.*